

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 19:44:02 ; Search time 1535.11 Seconds
(without alignments)
236.425 Million cell updates/sec

Title: US-09-802-518-1

Perfect score: 22

Sequence: 1 tgactgtgaacgttcgagatga 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_om.*

20: em_or.*

21: em_ov.*

22: em_pat.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_sts.*

27: em_sy.*

28: em_un.*

29: em_vi.*

30: em_htgo_hum.*

31: em_htgo_inv.*

32: em_htgo_rod.*

33: em_htg_hum.*

34: em_htg_inv.*

35: em_htg_rod.*

36: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	22	100.0	22	6	AX036945	Sequence
2	22	100.0	22	6	AX046993	Sequence
3	22	100.0	22	6	AX083675	Sequence
4	22	100.0	22	6	AX135650	Sequence
5	22	100.0	22	6	AX148636	Sequence
6	21	95.5	22	6	AX083681	Sequence
7	21	95.5	22	6	AX148642	Sequence
8	20.4	92.7	22	6	AR148608	Sequence
9	20.4	92.7	22	6	AX036946	Sequence
10	20.4	92.7	22	6	AX083676	Sequence
11	20.4	92.7	22	6	AX083678	Sequence
12	20.4	92.7	22	6	AX148637	Sequence
13	20.4	92.7	22	6	AX148639	Sequence
14	20.2	91.8	22	6	AX148643	Sequence
15	20	90.9	22	6	AX083682	Sequence
16	20	90.9	22	6	AX174913	Sequence
17	19.4	88.2	22	6	AX083680	Sequence
18	19.4	88.2	22	6	AX148641	Sequence
19	18.8	85.5	22	6	AR148607	Sequence
20	18.8	85.5	22	6	AR148609	Sequence
21	18.8	85.5	22	6	AR148616	Sequence
22	18.8	85.5	22	6	AX036944	Sequence
23	18.8	85.5	22	6	AX036952	Sequence
24	18.8	85.5	22	6	AX135651	Sequence
25	18.8	85.5	22	6	AX148644	Sequence
26	18.8	85.5	22	6	AX148645	Sequence
27	17.8	80.9	165337	2	AC027442	Homo sap
28	17.8	80.9	167237	9	AC007938	Homo sap
c 29	17.8	80.9	177371	2	AC092138	Homo sap
c 30	17.8	80.9	201214	2	AC074012	Homo sap
c 31	17.8	80.9	204992	2	AC024934	Homo sap
32	17.2	78.2	22	6	AR148610	Sequence
33	17.2	78.2	22	6	AX135652	Sequence
34	17.2	78.2	127	3	AGXH8	272011 A. gambiae s
c 35	17.2	78.2	552	8	RIC5INE04	D85048 Rice retrotr
c 36	17.2	78.2	12545	3	CEFA2G10	Z48230 Caenorhabd
c 37	17.2	78.2	85219	9	AC073520	Homo sap
38	17.2	78.2	110000	2	AC091528_2	Continuation (3 of
39	17.2	78.2	128216	9	HSDJ319M7	AL079341 Human DNA
40	17.2	78.2	169271	2	AC079336	AC079336 Homo sap
41	17.2	78.2	169277	2	AC078940	Homo sap
42	17.2	78.2	183827	9	AC005899	AC005899 Homo sap
43	17.2	78.2	185784	2	AC009719	Homo sap
44	17.2	78.2	265040	2	AC091363	AC091363 Rattus no
45	16.8	76.4	2040	3	AF115507	AF115507 Leishman

ALIGNMENTS

RESULT 1
AX036945
LOCUS AX036945 22 bp DNA
DEFINITION Sequence 2 from Patent FR2790955.
ACCESSION AX036945 PAT
VERSION AX036945.1 GI:11226373
SOURCE
synthetic construct.
synthetic construct
artificial sequence.
ORGANISM
1 (bases 1 to 22)
REFERENCE
Carpentier,A.
AUTHORS
Patent: FR 2790955-A 2 22-SEP-2000;
JOURNAL ASSIST PUBL HOPITAUX DE PARIS (FR)
FEATURES
Location/Qualifiers
source
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note=" oligodesoxynucleotide"
BASE COUNT 6 a 3 c 7 g 6 t

ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagatga 22
 |||||
 Db 1 TGACTGTGAACGTCGAGATGA 22

RESULT 2

AX046993 AX046993 22 bp DNA PAT 15-DEC-2000
 LOCUS Sequence 2 from Patent WO0067787.
 DEFINITION AX046993
 ACCESSION AX046993
 VERSION AX046993.1 GI:11876420
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1. .22
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="phosphorothioate-modified synthetic oligodeoxynucleotide"
 6 a 3 c 7 g 6 t

BASE COUNT

ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagatga 22
 |||||
 Db 1 TGACTGTGAACGTCGAGATGA 22

RESULT 3

AX083675 AX083675 22 bp DNA PAT 28-FEB-2001
 LOCUS Sequence 1 from Patent WO0112223.
 DEFINITION AX083675
 ACCESSION AX083675
 VERSION AX083675.1 GI:13185407
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1. .22
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Synthetic construct"
 6 a 3 c 7 g 6 t

BASE COUNT

ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.13;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagatga 22
 |||||
 Db 1 TGACTGTGAACGTCGAGATGA 22

RESULT 4

AX135650 AX135650 22 bp DNA PAT 29-MAY-2001
 LOCUS Sequence 21 from Patent WO0132877.
 DEFINITION AX135650
 ACCESSION AX135650
 VERSION AX135650.1 GI:14271920
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1. .22
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="CpG oligonucleotide"
 6 a 3 c 7 g 6 t

Query Match 100.0%; Score 22; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagatga 22
 |||||
 Db 1 TGACTGTGAACGTCGAGATGA 22

RESULT 5

AX148636 AX148636 22 bp DNA PAT 08-JUN-2001
 LOCUS Sequence 1 from Patent WO0135991.
 DEFINITION AX148636
 ACCESSION AX148636
 VERSION AX148636.1 GI:14347254
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1. .22
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="synthetic construct"
 6 a 3 c 7 g 6 t

Query Match 100.0%; Score 22; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagatga 22
 |||||
 Db 1 TGACTGTGAACGTCGAGATGA 22

RESULT 6
AX083681
LOCUS AX083681 22 bp DNA PAT 28-FEB-2001
DEFINITION Sequence 7 from Patent WO0112223.
ACCESSION AX083681
VERSION AX083681.1 GI:13185413
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS van Nest,G.
TITLE Methods of modulating an immune response using immunostimulatory s
sequences and compositions for use therein
JOURNAL Patent: WO 0112223-A 7 22-FEB-2001;
Dynamax Technologies Corporation (US)
FEATURES
source 1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
modified_base 11
/note="5-bromocytosine"
/mod_base=OTHER
BASE COUNT 6 a 2 c 7 g 6 t 1 others
ORIGIN
Query Match 95.5%; Score 21; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 0.49;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 tgactgtgaacgttcgagatga 22
||||| ||||| ||||| |||||
Db 1 TGACTGTGAANGTTCGAGATGA 22
RESULT 7
AX148642
LOCUS AX148642 22 bp DNA PAT 08-JUN-2001
DEFINITION Sequence 7 from Patent WO0135991.
ACCESSION AX148642
VERSION AX148642.1 GI:14347260
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS Tuck,S. and van Nest,G.
TITLE Immunomodulatory compositions containing an immunostimulatory
sequence linked to antigen and methods of use thereof
JOURNAL Patent: WO 0135991-A 7 25-MAY-2001;
Dynamax Technologies Corporation (US)
FEATURES
source 1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic construct"
modified_base 11
/note="5-bromocytosine"
/mod_base=OTHER
BASE COUNT 6 a 2 c 7 g 6 t 1 others
ORIGIN
Query Match 95.5%; Score 21; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 0.49;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 tgactgtgaacgttcgagatga 22
||||| ||||| ||||| |||||
Db 1 TGACTGTGAANGTTCGAGATGA 22

RESULT 8
AR148608
LOCUS AR148608 22 bp DNA PAT 08-AUG-2001
DEFINITION Sequence 2 from patent US 6225292.
ACCESSION AR148608
VERSION AR148608.1 GI:15112698
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 22)
AUTHORS Raz,E. and Roman,M.
TITLE Inhibitors of DNA immunostimulatory sequence activity
JOURNAL Patent: US 6225292-A 2 01-MAY-2001;
FEATURES
source 1. .22
/organism="unknown"
BASE COUNT 7 a 2 c 7 g 6 t
ORIGIN
Query Match 92.7%; Score 20.4; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 tgactgtgaacgttcgagatga 22
||||| ||||| ||||| |||||
Db 1 TGACTGTGAACGTTAGAGATGA 22
RESULT 9
AX036946
LOCUS AX036946 22 bp DNA PAT 16-NOV-2000
DEFINITION Sequence 3 from Patent FR2790955.
ACCESSION AX036946
VERSION AX036946.1 GI:11226374
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS Carpentier,A.
JOURNAL Patent: FR 2790955-A 3 22-SEP-2000;
ASSIST PUBL HOPITAUX DE PARIS (FR)
FEATURES
source 1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligodesoxynucleotide"
BASE COUNT 6 a 4 c 6 g 6 t
ORIGIN
Query Match 92.7%; Score 20.4; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 tgactgtgaacgttcgagatga 22
||||| ||||| ||||| |||||
Db 1 TGACTGTGAACGTTCCAGATGA 22
RESULT 10
AX083676
LOCUS AX083676 22 bp DNA PAT 28-FEB-2001
DEFINITION Sequence 2 from Patent WO0112223.
ACCESSION AX083676
VERSION AX083676.1 GI:13185408
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequence.

REFERENCE 1 (bases 1 to 22)
AUTHORS van Nest,G.
TITLE Methods of modulating an immune response using immunostimulatory s
sequences and compositions for use therein
JOURNAL Patent: WO 0112223-A 2 22-FEB-2001;
Dynavax Technologies Corporation (US)
FEATURES Location/Qualifiers
source
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic construct"
BASE COUNT 6 a 4 c 7 g 5 t
ORIGIN

Query Match 92.7%; Score 20.4; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
||||| ||||||| ||||||| |||||||
Db 1 TGACCGTGAACGTTCCGAGATGA 22

RESULT 11
AX083678 22 bp DNA PAT 28-FEB-2001
LOCUS
DEFINITION Sequence 4 from Patent WO0112223.
ACCESSION AX083678
VERSION AX083678.1 GI:13185410
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS van Nest,G.
TITLE Methods of modulating an immune response using immunostimulatory s
sequences and compositions for use therein
JOURNAL Patent: WO 0112223-A 4 22-FEB-2001;
Dynavax Technologies Corporation (US)
FEATURES Location/Qualifiers
source
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic construct"
BASE COUNT 6 a 4 c 6 g 6 t
ORIGIN

Query Match 92.7%; Score 20.4; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
||||| ||||||| ||||||| |||||||
Db 1 TGACGTGAACGTTCCAGATGA 22

RESULT 12
AX148637 22 bp DNA PAT 08-JUN-2001
LOCUS
DEFINITION Sequence 2 from Patent WO0135991.
ACCESSION AX148637
VERSION AX148637.1 GI:14347255
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Tuck,S. and van Nest,G.
TITLE Immunomodulatory compositions containing an immunostimulatory
sequence linked to antigen and methods of use thereof
JOURNAL Patent: WO 0135991-A 2 25-MAY-2001;

FEATURES Dynavax Technologies Corporation (US)
Location/Qualifiers
source
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic construct"
BASE COUNT 6 a 4 c 7 g 5 t
ORIGIN

Query Match 92.7%; Score 20.4; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
||||| ||||||| ||||||| |||||||
Db 1 TGACCGTGAACGTTCCGAGATGA 22

RESULT 13
AX148639 22 bp DNA PAT 08-JUN-2001
LOCUS
DEFINITION Sequence 4 from Patent WO0135991.
ACCESSION AX148639
VERSION AX148639.1 GI:14347257
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Tuck,S. and van Nest,G.
TITLE Immunomodulatory compositions containing an immunostimulatory
sequence linked to antigen and methods of use thereof
JOURNAL Patent: WO 0135991-A 4 25-MAY-2001;
Dynavax Technologies Corporation (US)
FEATURES Location/Qualifiers
source
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic construct"
BASE COUNT 6 a 4 c 6 g 6 t
ORIGIN

Query Match 92.7%; Score 20.4; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
||||| ||||||| ||||||| |||||||
Db 1 TGACTGTGAACGTTCCAGATGA 22

RESULT 14
AX148643 22 bp DNA PAT 08-JUN-2001
LOCUS
DEFINITION Sequence 8 from Patent WO0135991.
ACCESSION AX148643
VERSION AX148643.1 GI:14347261
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Tuck,S. and van Nest,G.
TITLE Immunomodulatory compositions containing an immunostimulatory
sequence linked to antigen and methods of use thereof
JOURNAL Patent: WO 0135991-A 8 25-MAY-2001;
Dynavax Technologies Corporation (US)
FEATURES Location/Qualifiers
source
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"

```
modified_base 11 /note="synthetic construct"
/modified_base 11 /note="5-bromocytosine"
BASE COUNT 6 a 1 c 7 g 6 t 2 others
ORIGIN
```

```
Query Match 91.8%; Score 20.2; DB 6; Length 22;
Best Local Similarity 90.9%; Pred. No. 1.4;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 1 tgactgtgaacgttcgagatga 22
Db 1 TGACTGTGAANGTTBGAGATGA 22
```

```
RESULT 15
AX083682 AX083682 22 bp DNA PAT 28-FEB-2001
DEFINITION Sequence 8 from Patent WO0112223.
ACCESSION AX083682
VERSION AX083682.1 GI:13185414
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS van Nest,G.
TITLE Methods of modulating an immune response using immunostimulatory s
JOURNAL sequences and compositions for use therein
Patent: WO 0112223-A 8 22-FEB-2001;
Dynamax Technologies Corporation (US)
FEATURES Location/Qualifiers
Source 1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
```

```
modified_base 11 /note="5-bromocytosine"
/modified_base 15 /note="5-bromocytosine"
BASE COUNT 6 a 1 c 7 g 6 t 2 others
ORIGIN
```

```
Query Match 90.9%; Score 20; DB 6; Length 22;
Best Local Similarity 90.9%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 tgactgtgaacgttcgagatga 22
Db 1 TGACTGTGAANGTTNGAGATGA 22
```

Search completed: March 11, 2002, 20:55:45
Job time: 4303 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 19:50:17 ; Search time 227.21 Seconds
(without alignments)
83.012 Million cell updates/sec

Title: US-09-802-518-1

Perfect score: 22

Sequence: 1 tgactgtgaacgttcgagatga 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101.*
1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	19	AAV32079
2	22	100.0	22	20	AAV36624
3	22	100.0	22	20	AAV80097
4	22	100.0	22	20	AAV80102
5	22	100.0	22	20	AAV80103
6	22	100.0	22	21	AAV80105
7	22	100.0	22	21	AAV80105
8	22	100.0	22	21	AAV80105
9	22	100.0	22	21	AAV80105
10	22	100.0	22	21	AAV80105
11	22	100.0	22	21	AAV80105

12	22	100.0	22	21	AAV38072	Immunostimulatory
13	22	100.0	22	21	AAV38072	Immunostimulatory
14	22	100.0	22	22	AAH42533	Phosphorothioate b
15	22	100.0	22	22	AAH73439	Immunomodulatory n
16	22	100.0	22	22	AAH44109	Immunostimulatory
17	22	100.0	22	22	AAH41573	5' terminal NH2 gr
18	22	100.0	22	22	AAH20403	CpG motif containi
19	22	100.0	22	22	AAH77040	Immunomodulatory b
20	22	100.0	22	22	AAH29800	Cholera toxin immu
21	22	100.0	22	22	AAH82107	Oligonucleotide OD
22	22	100.0	22	22	AAH92377	CG motif and CFA c
23	21.2	96.4	22	22	AAH77046	Immunostimulatory
24	21	95.5	22	21	AAH55880	Immunomodulatory o
25	21	95.5	22	22	AAH41579	Immunostimulatory
26	20.4	92.7	22	20	AAV80105	Oligo used in expe
27	20.4	92.7	22	20	AAV80096	Immunomodulatory o
28	20.4	92.7	22	20	AAV80099	Immunomodulatory o
29	20.4	92.7	22	20	AAV80101	Immunomodulatory o
30	20.4	92.7	22	21	AAH96254	Sequence of a stab
31	20.4	92.7	22	21	AAH38066	Immunostimulatory
32	20.4	92.7	22	21	AAH38066	Immunostimulatory
33	20.4	92.7	22	21	AAH38070	Immunostimulatory
34	20.4	92.7	22	22	AAH42534	Phosphorothioate b
35	20.4	92.7	22	22	AAH73440	Immunomodulatory n
36	20.4	92.7	22	22	AAH41574	Immunostimulatory
37	20.4	92.7	22	22	AAH41576	Immunostimulatory
38	20.4	92.7	22	22	AAH77041	Immunostimulatory
39	20.4	92.7	22	22	AAH77043	Immunostimulatory
40	20.4	92.7	22	22	AAH77047	Immunostimulatory
41	20.2	91.8	22	22	AAH41580	Immunostimulatory
42	20	90.9	22	21	AAH55881	Immunomodulatory o
43	19.6	89.1	22	22	AAH77045	Immunostimulatory
44	19.4	88.2	22	21	AAH55877	Immunomodulatory o
45	19.4	88.2	22	22	AAH41578	Immunostimulatory

ALIGNMENTS

RESULT 1
AAV32079 standard; DNA; 22 bp.
ID AAV32079 standard; DNA; 22 bp.
XX
AC AAV32079;
XX
DT 09-SEP-1998 (first entry)
XX
DE Nucleotide sequence of DY1018.
XX
DY1018; beta-gal; ISS-PN/IMM; antigen; immune response; antibody;
KW immunisation; anaphylaxis; IgE; retinopathies; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..22
FT /*tag= a
FT /note= "phosphothioate backbone"
XX
PN W09816247-A1.
XX
PD 23-APR-1998.
XX
PF 09-OCT-1997; 97WO-US19004.
XX
PR 11-OCT-1996; 96US-0028118.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Carson DA, Raz E, Roman M;
XX WPI; 1998-361028/23.
XX

PT New immunomodulatory compositions - comprising an antigen conjugated
 PT to a polynucleotide that contains an immunostimulatory sequence
 XX
 PS Example 1; Page 36; 69pp; English.

XX This is the nucleotide sequence of DY1018, which is conjugated to
 CC beta-gal to form ISS-PN/IMM, comprising an immunomodulatory molecule
 CC (IMM) which comprises an antigen conjugated to a polynucleotide
 CC (PN) that contains at least one immunostimulatory nucleotide sequence
 CC (ISS). The conjugate synergistically boost the magnitude of the host
 CC immune response against an antigen to a level greater than the host
 CC immune response to either the IMM, antigen or ISS-PN alone. These
 CC responses to ISS-PN/IMM conjugates are particularly acute during
 CC the important early phase of the host immune response to an antigen.
 CC The ISS-PN/IMM conjugates boost both humoral (antibody) and cellular
 CC (Th1 type) immune responses of the host. Thus, use of the method to
 CC boost the immune responsiveness of a host to subsequent challenge by a
 CC sensitising antigen without immunisation avoids the risk of
 CC Th2-mediated, immunisation-induced anaphylaxis by suppressing IgE
 CC production in response to the antigen challenge. The conjugates can
 CC also be used to combat pathogenic infection and to stimulate
 CC therapeutic angiogenesis to treat conditions in which localised blood
 CC flow plays a significant etiological role, e.g. retinopathies.
 XX
 SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 19; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgactgtgaacgttcgagatga 22
 |||||
 Db 1 tgactgtgaacgttcgagatga 22

RESULT 2
 AAX36624
 ID AAX36624 standard; DNA; 22 BP.
 XX
 AC AAX36624;
 XX
 DT 09-JUL-1999 (first entry)
 XX
 DE ISS-ODN DY1018 nucleotide sequence.
 XX
 KW Antigen-stimulated inflammation; immunostimulatory oligonucleotide;
 KW granulocyte-mediated tissue inflammation; Th2 type immune response;
 KW immune responsiveness modulation; idiopathic hypereosinophilic syndrome;
 KW cutaneous basophil hypersensitivity; ISS-ODN; asthma; nasal polyposis;
 KW allergic rhinitis; atopic dermatitis; allergic conjunctivitis;
 KW eosinophilic fasciitis; therapy; ss.
 XX
 OS Synthetic.
 XX
 PN WO9911275-A2.
 XX
 PD 11-MAR-1999.
 XX
 PF 04-SEP-1998; 98WO-US18382.
 XX
 PR 05-SEP-1997; 97US-0927120.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Ray E;
 XX
 DR WPI; 1999-312404/26.
 XX
 PT Reducing antigen-stimulated granulocyte-mediated inflammation
 XX
 PS Example 2; Page 30; 69pp; English.
 XX

CC This is the ISS-ODN DY1018 nucleotide sequence.
 CC The invention relates to a method for preventing or reducing
 CC antigen-stimulated, granulocyte-mediated tissue inflammation in a
 CC by administering an immunostimulatory oligonucleotide (ISS-ODN), where
 CC (a) reduction in, or the absence of, a Th2 type immune response is
 CC measured; or (b) there is a reduction or absence of other clinical signs
 CC of inflammation in the host after antigen challenge. The method is used
 CC to reduce or suppress granulocyte-mediated inflammation in a host tissue,
 CC and to modulate the host's immune responsiveness to an antigen,
 CC particularly where the subject suffers from asthma, nasal polyposis,
 CC allergic rhinitis, atopic dermatitis, allergic conjunctivitis, or
 CC eosinophilic fasciitis, idiopathic hypereosinophilic syndrome, or
 CC cutaneous basophil hypersensitivity. Unlike prior art treatment by
 CC antigen immunisation, the method is an antigen-independent method,
 CC and avoids host production of both interleukin-4 (IL-4), which carries
 CC risk of anaphylaxis, and IL-5 which actually encourages granulocyte
 CC adhesion to endothelia.
 XX
 SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 20;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgactgtgaacgttcgagatga 22
 |||||
 Db 1 tgactgtgaacgttcgagatga 22

RESULT 3
 AAV80097
 ID AAV80097 standard; DNA; 22 BP.

XX
 AC AAV80097;
 XX
 DT 12-MAR-1999 (first entry)
 XX
 DE Immunomodulatory oligo comprising an ISS sequence.
 XX
 KW Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
 KW ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
 KW human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
 KW B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
 XX
 OS Synthetic.
 XX
 PN WO9855495-A2.
 XX
 PD 10-DEC-1998.
 XX
 PF 05-JUN-1998; 98WO-US11578.
 XX
 PR 06-JUN-1997; 97US-0048793.
 XX
 PA (DYNA-) DYNAVAX TECHNOLOGIES CORP.
 XX
 PI Dina D, Roman M, Schwartz D;
 XX
 WPI; 1999-059898/05.
 XX
 PT Immunostimulatory oligonucleotides regulate the immune system - and
 PT contain an immune-stimulating octanucleotide sequence; for treating
 PT cancer, allergic and infectious diseases
 XX
 PS Claim 5; Page 29; 63pp; English.
 XX
 CC The invention relates to immunomodulatory oligonucleotides that comprise
 CC at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
 CC sequences are selected from the group consisting of AACGTTC, AACGTTCG,
 CC GACGTTC, and GACGTTCG. The immunomodulatory sequences are used to treat
 CC patients needing immune regulation, such as those suffering from cancer,
 CC an allergic disease and asthma. They are also used to prevent infectious

CC diseases such as influenza, herpes, hepatitis B, human immunodeficiency
 CC and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
 CC Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
 CC Schistosoma. The immunomodulatory sequences are used to screen for human
 CC immunostimulatory activity by incubating macrophage cells and the
 CC oligonucleotide; and determining the relative amount of Thi-biased
 CC cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
 CC specific claimed examples of such immunomodulatory oligonucleotides.
 XX
 SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 20; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
 |||||
 Db 1 tgactgtgaacgttcgagatga 22

RESULT 4
 AAV80102
 ID AAV80102 standard; DNA: 22 BP.
 XX
 AC AAV80102;
 XX
 DT 12-MAR-1999 (first entry)
 XX
 DE Immunomodulatory oligo comprising an ISS sequence.
 XX
 KW Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
 KW ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
 KW human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
 KW B. pertussis; malaria; plasmodia; leishmania; trypanosoma; schistosoma.
 XX
 OS Synthetic.

Key Location/Qualifiers
 modified_base 11
 /*tag= a
 /note= "5-bromocytosine"

WO9855495-A2.
 10-DEC-1998.
 05-JUN-1998; 98WO-US11578.
 06-JUN-1997; 97US-0048793.
 (DYNA-) DYNAVAX TECHNOLOGIES CORP.

Dina D, Roman M, Schwartz D;
 WPI; 1999-059898/05.

Immunostimulatory oligonucleotides regulate the immune system - and
 contain an immune-stimulating octanucleotide sequence; for treating
 cancer, allergic and infectious diseases

Claim 23; Page 30; 63pp; English.

CC The invention relates to immunomodulatory oligonucleotides that comprise
 CC at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
 CC sequences are selected from the group consisting of AACGTTCC, AACGTTCCG,
 CC GAGGTTCC, and GAGGTTCCG. The immunomodulatory sequences are used to treat
 CC patients needing immune regulation, such as those suffering from cancer,
 CC an allergic disease and asthma. They are also used to prevent infectious
 CC diseases such as influenza, herpes, hepatitis B, human immunodeficiency
 CC and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
 CC Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
 CC Schistosoma. The immunomodulatory sequences are used to screen for human

CC immunostimulatory activity by incubating macrophage cells and the
 CC oligonucleotide; and determining the relative amount of Thi-biased
 CC cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
 CC specific claimed examples of such immunomodulatory oligonucleotides.
 XX
 SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 20; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
 |||||
 Db 1 tgactgtgaacgttcgagatga 22

RESULT 5
 AAV80103
 ID AAV80103 standard; DNA: 22 BP.
 XX
 AC AAV80103;
 XX
 DT 12-MAR-1999 (first entry)
 XX
 DE Immunomodulatory oligo comprising an ISS sequence.
 XX
 KW Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
 KW ISS: cancer; allergy; asthma; hepatitis B infection; papillomavirus;
 KW human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
 KW B. pertussis; malaria; plasmodia; leishmania; trypanosoma; schistosoma.
 XX
 OS Synthetic.

Key Location/Qualifiers
 modified_base 11
 /*tag= a
 /note= "5-bromocytosine"

WO9855495-A2.
 10-DEC-1998.
 05-JUN-1998; 98WO-US11578.
 06-JUN-1997; 97US-0048793.
 (DYNA-) DYNAVAX TECHNOLOGIES CORP.

Dina D, Roman M, Schwartz D;
 WPI; 1999-059898/05.

Immunostimulatory oligonucleotides regulate the immune system - and
 contain an immune-stimulating octanucleotide sequence; for treating
 cancer, allergic and infectious diseases

Claim 24; Page 30; 63pp; English.

CC The invention relates to immunomodulatory oligonucleotides that comprise
 CC at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
 CC sequences are selected from the group consisting of AACGTTCC, AACGTTCCG,
 CC GAGGTTCC, and GAGGTTCCG. The immunomodulatory sequences are used to treat
 CC patients needing immune regulation, such as those suffering from cancer,
 CC an allergic disease and asthma. They are also used to prevent infectious
 CC diseases such as influenza, herpes, hepatitis B, human immunodeficiency
 CC and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
 CC Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
 CC Schistosoma. The immunomodulatory sequences are used to screen for human
 CC immunostimulatory activity by incubating macrophage cells and the
 CC oligonucleotide; and determining the relative amount of Thi-biased
 CC cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
 CC specific claimed examples of such immunomodulatory oligonucleotides.

XX SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 20; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 tgactgtgaacgttcgagatga 22

RESULT 6

AAAC64051
ID AAC64051 standard; DNA; 22 BP.

XX AC AAC64051;

XX DT 15-FEB-2001 (first entry)

DE Immunostimulatory CpG phosphorothioate oligodeoxynucleotide.

XX CpG oligodeoxynucleotide; phosphorothioate; immunostimulatory; ISS ODN;
KW enhanced antigen presentation; antigen-presenting cell; APC;
KW T-cell activation; tumour cell; tumour antigen; cancer immunotherapy;
KW vaccine; ss.

XX OS Synthetic.

XX WO2000062787-A1.

XX PD 26-OCT-2000.

XX PF 11-APR-2000; 2000WO-US09664.

XX PR 15-APR-1999; 99US-0292278.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Raz E, Martin-Orozco E;

XX WPI; 2000-679548/66.

XX Enhancing antigen-presentation capabilities of T-cells for cancer
PT immunotherapy, by contacting cells with an immunostimulatory
PT oligonucleotide

XX Example I; Page 18; 42pp; English.

XX The invention relates to a method of inducing activation of T-cells
CC to respond to an antigen, comprising contacting antigen-presenting cells
CC (APC) with an immunostimulatory oligodeoxynucleotide (ISS-ODN). The APCs
CC thus treated have enhanced antigen presenting capabilities compared to
CC antigen-activated APCs. APCs with enhanced antigen-presentation
CC capabilities then present the antigen to T-cells. The method is useful
CC for cancer immunotherapy. The ISS-ODN is used to enhance the tumour
CC antigen presenting capacity of tumour cells, thereby inducing T-cell
CC activation, and is therefore useful for treating tumours. Additionally,
CC tumour cells treated with an ISS-ODN ex vivo are useful as vaccines.
CC ISS-ODN treated APCs are induced to take up antigen through upregulation
CC of Fc-receptor expression, to present antigen through upregulation of
CC major histocompatibility complex (MHC) Class I and II expression and
CC CD1d expression, to produce co-stimulatory factors (B7 and CD40), to
CC provide cell-to-cell adhesion through upregulation of intercellular
CC adhesion molecule (ICAM) expression, and to increase Th1 stimulatory
CC cytokine production, all at levels greater than that achieved through
CC contact of APC with antigen alone. The present sequence represents
CC a phosphorothioate CpG ISS-ODN used in the exemplifications of the
XX invention.

XX SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 tgactgtgaacgttcgagatga 22

RESULT 7

AAA96253
ID AAA96253 standard; DNA; 22 BP.

XX AC AAA96253;

XX DT 08-FEB-2001 (first entry)

XX Sequence of a stabilised oligonucleotide with antitumour activity.

XX Antitumour; immunostimulatory oligonucleotide; tumour; anaplasia;
KW glioblastoma; medullablastoma; neuroblastoma; carcinoma; ss.
XX Synthetic.

XX WO200056342-A2.

XX PD 28-SEP-2000.

XX PF 17-MAR-2000; 2000WO-FR00676.

XX PR 19-MAR-1999; 99FR-0003433.

XX PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

XX (INRM) INST NAT SANTE & RECH MEDICALE.

XX PI Carpentier A;

XX WPI; 2000-602192/57.

XX Use of stabilized oligonucleotides as antitumor agents, particularly
PT against nervous system tumors, have optimal activity and are not toxic

XX Example 2; Page 16; 57pp; French.

XX The present sequence represents a stabilised oligonucleotide which has
CC antitumour activity. The oligonucleotide comprises an octamer motif
CC of the type 5'-purine-purine-cg-pyrimidine-pyrimidine-x-x-3', where
CC the pair x-x is AT, AA, CT or TT. The oligonucleotides are
CC immunostimulatory, and are not toxic. They may be adapted for use in
CC animals or humans. The stabilised oligonucleotides are used for
CC treating tumours, of any type and any degree of anaplasia, particularly
CC human tumours in the peripheral or central nervous systems, specifically
CC glioblastomas, medullablastomas, neuroblastomas, melanomas or carcinomas.
XX Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 tgactgtgaacgttcgagatga 22

RESULT 8

AAA90458
ID AAA90458 standard; DNA; 22 BP.

XX AC AAA90458;

XX DT 10-JAN-2001 (first entry)
XX DE CpG adjuvant oligonucleotide, SEQ ID NO:19.
XX KW CpG oligonucleotide; CpG motif; adjuvant; microdroplet emulsion;
XX KW microemulsion; adsorbent microparticle; vaccine; Th1 immune response;
XX KW viral infection; bacterial infection; parasitic infection; HCV; HBV;
XX KW hepatitis C virus; hepatitis B virus; herpes simplex virus; HSV; HIV;
XX KW human immunodeficiency virus; cytomegalovirus; CMV; influenza virus;
XX KW rabies virus; cholera; diphtheria; tetanus; pertussis;
XX KW Helicobacter pylori; Haemophilus influenzae; malaria; ss.
XX OS Synthetic.
XX XX
XX XX
XX PN WO200050006-A2.
XX PD 31-AUG-2000.
XX PF 09-FEB-2000; 2000WO-US03331.
XX PR 26-FEB-1999; 99US-0121858.
XX PR 29-JUL-1999; 99US-0146391.
XX PR 28-OCT-1999; 99US-0161997.
XX XX
XX PA (CHIR) CHIRON CORP.
XX XX
XX PI O'Hagan D, Ott GS, Donnelly J, Kazzaz J, Ugozzoli M, Singh M;
XX PI Barackman J;
XX XX
XX DR WPI; 2000-587123/55.
XX XX
XX PT Microemulsion having an adsorbent surface comprising a microdroplet
XX PT emulsion consisting of a metabolizable oil and an emulsifying agent
XX PT which is a detergent, useful as a vaccine to treat bacterial, viral,
XX PT and parasitic infection
XX XX
XX PS Claim 17; Page 40; 95pp; English.
XX XX
XX CC The invention relates to a microdroplet emulsion (microemulsion) with an
XX CC adsorbent surface, and which comprises a metabolizable oil and an
XX CC emulsifying agent (a detergent). It also relates to a composition
XX CC comprising the microemulsion and a microparticle with an adsorbent
XX CC surface, where the microparticle comprises a polymer selected from a
XX CC poly(alpha-hydroxy acid), a polyhydroxy butyric acid, a
XX CC polycaprolactone, a polyorthoester, a polyanhydride, and a
XX CC polycyanacrylate, and a second detergent. The surface of the
XX CC microparticles efficiently adsorb biologically active macromolecules such
XX CC as DNA, polypeptides, antigens, hormones, pharmaceuticals, enzymes,
XX CC mediators of transcription or translation, metabolic intermediates and
XX CC adjuvants. Additionally, a second biologically active molecule may be
XX CC encapsulated within the microparticle. The microemulsion can be used in
XX CC methods of immunising a host animal, particularly a human, against a
XX CC viral, bacterial or parasitic infection, and in methods of increasing a
XX CC Th1 immune response. The microemulsions (having the appropriate antigens
XX CC adsorbed) may be particularly used as vaccines for hepatitis C virus
XX CC (HCV), hepatitis B virus (HBV), herpes simplex virus (HSV), human
XX CC immunodeficiency virus (HIV), cytomegalovirus (CMV), influenza virus, and
XX CC rabies virus; the bacteria which cause cholera, diphtheria, tetanus and
XX CC pertussis; Helicobacter pylori and Haemophilus influenzae; and
XX CC malaria-causing parasites. Sequences AAA90447-A90467 represent Th1
XX CC lymphocyte stimulating oligonucleotides containing at least one CpG motif
XX CC which are claimed for use as adjuvants in the compositions of the
XX CC invention.
XX XX
XX SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 tgactgtgaacgttcgagatga 22

Db 1 tgactgtgaacgttcgagatga 22
RESULT 9
AAAL4467
ID AAA14467 standard; DNA; 22 BP.
XX AC AAAL4467;
XX DT 21-AUG-2000 (first entry)
XX XX
XX DE Immunostimulatory oligonucleotide (ISS-ODN) DY1018.
XX XX
XX KW Immunostimulatory oligonucleotide; adjuvant; mucosal immunity;
XX KW secretory immunoglobulin A production; sIgA; Th1 phenotype; ds.
XX OS Synthetic.
XX XX
XX PN WO200020039-A1.
XX PD 13-APR-2000.
XX PF 15-SEP-1999; 99WO-US21203.
XX PR 05-OCT-1998; 98US-0167039.
XX XX
XX PA (REGC) UNIV CALIFORNIA.
XX PI Raz E, Horner AA, Carson DA;
XX DR WPI; 2000-303647/26.
XX XX
XX PT Immunostimulatory oligonucleotide adjuvant induces mucosal immunity to
XX PT an antigen in a mammalian host through production of secretory
XX PT immunoglobulin A
XX XX
XX PS Claim 8; Page 21; 64pp; English.
XX XX
XX CC The invention relates to a method of inducing mucosal immunity to an
XX CC antigen in a mammalian host, including the production of secretory
XX CC immunoglobulin A (sIgA). Immune protection in the mucosa (the principal
XX CC site of entry of most foreign antigens) is mediated by mucosa-associated
XX CC lymphoid tissue, epithelial and distinct B-cell, T-cell and accessory
XX CC cell sub-populations. The primary immune response which characterises
XX CC the induction of mucosal immunity to an antigen is sIgA production by
XX CC activated B-cells. The method comprises introducing an immunostimulatory
XX CC oligonucleotide (ISS-ODN) and the antigen into host mucosa, where the
XX CC ISS-ODN includes a core nucleotide sequence. The core nucleotide
XX CC sequence is 5'-Purine-Purine-C-G-pyrimidine-pyrimidine-3', specific
XX CC examples of which are AACGTT, AGCGTC and GACGTT (SEQ ID NOS 1-3). A
XX CC specific example of an ISS-ODN is DY1018 (AAA14467). The ISS-ODN is used
XX CC as an adjuvant with an antigen for stimulating mucosal immunity. The
XX CC level of sIgA production induced in the host is at least 3 times the
XX CC magnitude of sIgA production achievable in response to introduction of
XX CC antigen alone into the mucosal tissue and is equivalent or greater than
XX CC the magnitude of sIgA production achievable in response to introduction
XX CC of the antigen and cholera toxin adjuvant into the mucosal tissue. The
XX CC host immune response is stimulated to antigen specific IgA production,
XX CC biased towards the Th1 phenotype while antigen-induced IgE production is
XX CC avoided. The adjuvant has little or no known toxicity in mammals and its
XX CC efficacy is comparable to that of cholera toxin which is used as a
XX CC mucosal adjuvant. The present sequence represents the immunostimulatory
XX CC oligonucleotide DY1018.
XX XX
XX SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 tgactgtgaacgttcgagatga 22

```

Db      1 tgactgtgaacgttcgagatga 22
      |||||||
RESULT 10
AAA38065
ID AAA38065 standard; DNA; 22 BP.
XX
AC AAA38065;
XX
DT 24-AUG-2000 (first entry)
XX
DE Immunostimulatory sequence (ISS) #1.
XX
KW Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
KW gp120; human immunodeficiency virus; HIV; Immune response; Infection;
KW development; ss.
XX
OS Synthetic.
XX
PN WO200021556-A1.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23677.
XX
PR 09-OCT-1998; 98US-0103733.
PR 07-OCT-1999; 99US-0415186.
XX
PA (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI Tighe H, Raz E, Schwartz D, Takabayashi K;
XX WPI; 2000-317846/27.
XX
PT Anti-HIV composition comprises immunostimulatory polynucleotides and
PT HIV glycoprotein gp120 useful for modulating, stimulating an immune
PT response against HIV in an HIV infected individual
XX
PS Claim 3; Page 16; 65pp; English.
XX
CC The present invention relates to an immunostimulatory composition
CC comprising a human immunodeficiency virus (HIV) antigen, and an
CC immunomodulatory polynucleotide comprising an immunostimulatory sequence
CC (ISS). This sequence represents an ISS that can be used in the
CC composition. An immunostimulatory composition which comprises a gp120
CC conjugated to it and not conjugated, is used for modulating or
CC stimulating a specific immune response against gp120 in an individual by
CC producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It
CC is also used for suppressing or delaying development of HIV infection in
CC an individual infected with HIV or an individual at risk of infection
CC with HIV, respectively. It is also used for treating an individual
CC infected with HIV in need of immune modulation.
XX
SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
      |||||||
Db 1 tgactgtgaacgttcgagatga 22

RESULT 11
AAA38071
ID AAA38071 standard; DNA; 22 BP.
XX
AC AAA38071;
XX
DT 24-AUG-2000 (first entry)
XX
DE Immunostimulatory sequence (ISS) #7.
XX

```

```

DT 24-AUG-2000 (first entry)
XX
DE Immunostimulatory sequence (ISS) #7.
XX
KW Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
KW gp120; human immunodeficiency virus; HIV; Immune response; Infection;
KW development; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base II /*tag= a
FT /mod_base= OTHER
FT /note= "5-Bromocytosine"
XX
PN WO200021556-A1.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23677.
XX
PR 09-OCT-1998; 98US-0103733.
PR 07-OCT-1999; 99US-0415186.
XX
PA (DYNA-) DYNAVAX TECHNOLOGIES CORP.
XX
PI Tighe H, Raz E, Schwartz D, Takabayashi K;
XX WPI; 2000-317846/27.
XX
PT Anti-HIV composition comprises immunostimulatory polynucleotides and
PT HIV glycoprotein gp120 useful for modulating, stimulating an immune
PT response against HIV in an HIV infected individual
XX
PS Disclosure; Page 17; 65pp; English.
XX
CC The present invention relates to an immunostimulatory composition
CC comprising a human immunodeficiency virus (HIV) antigen, and an
CC immunomodulatory polynucleotide comprising an immunostimulatory sequence
CC (ISS). This sequence represents an ISS that can be used in the
CC composition. An immunostimulatory composition which comprises a gp120
CC conjugated to it and not conjugated, is used for modulating or
CC stimulating a specific immune response against gp120 in an individual by
CC producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It
CC is also used for suppressing or delaying development of HIV infection in
CC an individual infected with HIV or an individual at risk of infection
CC with HIV, respectively. It is also used for treating an individual
CC infected with HIV in need of immune modulation.
XX
SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
      |||||||
Db 1 tgactgtgaacgttcgagatga 22

RESULT 12
AAA38072
ID AAA38072 standard; DNA; 22 BP.
XX
AC AAA38072;
XX
DT 24-AUG-2000 (first entry)
XX
DE Immunostimulatory sequence (ISS) #7.
XX

```

KW Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
 KW gp120; human immunodeficiency virus; HIV; immune response; infection;
 KW development; ss.

XX Synthetic.

XX Key Location/Qualifiers
 FH modified_base 11
 FT /*tag= a
 FT /mod_base= OTHER
 FT /note= "5-Bromocytosine"
 FT modified_base 15
 FT /*tag= b
 FT /mod_base= OTHER
 FT /note= "5-Bromocytosine"

XX WO200021556-A1.

XX 20-APR-2000.

XX 08-OCT-1999; 99WO-US23677.

XX 09-OCT-1998; 98US-0103733.

XX 07-OCT-1999; 99US-0415186.

XX (DYNA-) DYNAVAX TECHNOLOGIES CORP.

XX Tighe H, Raz E, Schwartz D, Takabayashi K;

XX WPI: 2000-317846/27.

XX Anti-HIV composition comprises immunostimulatory polynucleotides and
 PT HIV glycoprotein gp120 useful for modulating, stimulating an immune
 PT response against HIV in an HIV infected individual -

XX Disclosure; Page 17: 65pp; English.

XX The present invention relates to an immunostimulatory composition
 CC comprising a human immunodeficiency virus (HIV) antigen, and an
 CC immunomodulatory polynucleotide comprising an immunostimulatory sequence
 CC (ISS). This sequence represents an ISS that can be used in the
 CC composition. An immunostimulatory composition which comprises a gp120
 CC conjugated to an immunomodulatory polynucleotide, or is proximately
 CC associated to it and not conjugated, is used for modulating or
 CC stimulating a specific immune response against gp120 in an individual by
 CC producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It
 CC is also used for suppressing or delaying development of HIV infection in
 CC an individual infected with HIV or an individual at risk of infection
 CC with HIV, respectively. It is also used for treating an individual
 CC infected with HIV in need of immune modulation.

XX Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
 Best Local Similarity 100.0%; Pred. NO. 0.034;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
 |||||
 Db 1 tgactgtgaacgttcgagatga 22

RESULT 13

AAZ5876

ID AAZ5876 standard; DNA; 22 BP.

XX AAZ5876;

XX 10-APR-2000 (first entry)

XX Immunomodulatory oligonucleotide SEQ ID NO: 1.

XX

KW Immunomodulation; immunostimulatory sequence; adjuvant;
 KW Th1 immune response; cytotoxic T-cell; cytokine; cancer; allergy;
 KW asthma; immunosuppression; ss.

OS Mus musculus.
 OS Synthetic.

XX Key Location/Qualifiers
 FH modified_base 1..22
 FT /*tag= a
 FT /note= "Phosphorothioate linkages"
 FT misc_feature 9..16
 FT /*tag= b
 FT /note= "Immunostimulatory sequence (ISS)"

XX WO962923-A2.

XX 09-DEC-1999.

XX 04-JUN-1999; 99WO-US12538.

XX 05-JUN-1998; 98US-0088310.

XX 01-JUN-1999; 99US-0324191.

XX (DYNA-) DYNAVAX TECHNOLOGIES CORP.

XX Schwartz D;

XX WPI: 2000-105687/09.

XX Novel immunomodulatory oligonucleotide used to induce a Th1-type immune
 PT response, e.g. to tumor antigens -

XX Example 1; Page 35; 54pp; English.

XX Sequences AAZ5876-255877 and AAZ5880-255886 represent immunomodulatory
 CC oligonucleotides comprising an immunostimulatory sequence (ISS, e.g.,
 CC AACGTC, AACGTT, AGCGTC, AGCGTT, AGCGTC, GACGTC, GACGTT, GGCCTT,
 CC AACGTTCC and GACGTTCC). The invention relates to oligonucleotides
 CC comprising one or more ISSs, where the ISS comprises at least
 CC one modified cytosine with an electron-withdrawing moiety at
 CC position C-5 or C-6 of the base. Sequences AAZ5877 and AAZ5880-255886
 CC contain ISSs comprising at least one bromocytosine, whereas sequence
 CC AAZ5876 contains an unmodified ISS. The immunomodulatory
 CC oligonucleotides have an adjuvant-like effect; when formulated with an
 CC antigen, the oligonucleotides stimulate production of Th1-type cytokines,
 CC and induce a Th1-type immune response (activation of cytotoxic T cells),
 CC while simultaneously downregulating the Th2-type response. The Th1
 CC response is particularly effective for control of viruses and
 CC intracellular parasites. The immunomodulatory oligonucleotides are used,
 CC particularly when formulated with an antigen or a facilitator, for
 CC modulating immune responses. Such compositions may be used in tumour
 CC therapy, in treatment of allergy (including asthma), for inducing a
 CC vigorous cellular response (against a virus, bacterium, fungus or
 CC protozoan), and also in contraceptive vaccines based on sperm antigens.

XX Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
 Best Local Similarity 100.0%; Pred. NO. 0.034;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22

|||||

Db 1 tgactgtgaacgttcgagatga 22

RESULT 14

AAH42533

ID AAH42533 standard; DNA; 22 BP.

XX AAH42533;

```

XX 01-OCT-2001 (first entry)
XX Phosphorothioate beta-gal/immunostimulatory oligonucleotide.
DE
XX
XX Anaphylactic hypersensitivity; immunomodulatory nucleic acid; vaccine;
KW anaphylaxis-associated symptom; IgE; histamine; phosphorothioate; ss.
XX
XX Synthetic.
XX
XX WO200145750-A1.
XX
XX 28-JUN-2001.
XX
XX 20-DEC-2000; 2000WO-US35064.
XX
XX 21-DEC-1999; 99US-0171830.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Raz E, Horner AA;
XX
XX WPI; 2001-475812/51.
XX
XX Reducing risk of anaphylactic hypersensitivity response to an allergen
PT in a subject, by administering an immunomodulating nucleic acid
PT molecule comprising a specific sequence -
XX
XX Example 1; Page 22; 39pp; English.
XX
XX The specification describes a method for reducing a symptom associated
CC with anaphylactic hypersensitivity or risk of anaphylactic response in
CC a subject. The method comprises administering to an individual a
CC nucleic acid molecule comprising an immunomodulatory nucleic acid
CC molecule (INA) comprising the sequence 5'-C-G-3' to reduce
CC anaphylaxis-associated symptom. The method is useful for reducing a
CC symptom associated with anaphylactic hypersensitivity, including
CC elevated IgE level, elevated histamine level, constriction of the
CC airways and difficult breathing which can lead to anaphylactic reaction
CC or anaphylactic shock, thereby reducing the risk of death. The present
CC sequence represents a beta-gal/immunostimulatory sequence, which was
CC used as a vaccine to protect against the development of anaphylactic
CC hypersensitivity.
XX
XX Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;
XX
XX Query Match 100.0%; Score 22; DB 22; Length 22;
XX Best Local Similarity 100.0%; Pred. No. 0.034;
XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 tgactgtgaacgttcgagatga 22
Qy |||||||
XX 1 tgactgtgaacgttcgagatga 22
Db |||||||

RESULT 15
AAH73439
ID AAH73439 standard; DNA; 22 BP.
XX
XX AAH73439;
AC
XX
XX 01-OCT-2001 (first entry)
XX
XX Immunomodulatory nucleic acid.
XX
XX G3PDH gene; immunomodulatory oligonucleotide; infection; mycobacterium;
KW intracellular pathogen; anti-pathogenic; ss.
XX
XX Unidentified.
XX
XX OS
XX
XX WO200155341-A2.
XX
XX

```

```

PD 02-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US03029.
XX
XX 31-JAN-2000; 2000US-0179353.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Raz E, Kornbluth R, Catanzaro A, Hayashi T, Carson DA;
XX
XX WPI; 2001-483234/52.
XX
XX Treating infection of intracellular pathogen e.g., Mycobacterium, in a
XX subject, involves administering immunomodulatory nucleic acid molecule
XX to inhibit intracellular replication of intracellular pathogen -
XX
XX Examples; Page 26; 54pp; English.
XX
XX The present invention describes a method of treating an infection caused
XX by an intracellular pathogen, involving administering to the patient an
XX immunomodulatory nucleic acid and an anti-pathogenic agent. This is
XX particularly useful in the treatment of mycobacterial infections. The
XX present sequence is an immunomodulatory nucleic acid described in the
XX exemplification of the invention.
XX
XX Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;
XX
XX Query Match 100.0%; Score 22; DB 22; Length 22;
XX Best Local Similarity 100.0%; Pred. No. 0.034;
XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 tgactgtgaacgttcgagatga 22
Qy |||||||
XX 1 tgactgtgaacgttcgagatga 22
Db |||||||

Search completed: March 11, 2002, 20:59:44
Job time: 4167 sec

```

> O <
O | O Intelligence
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

-- Outline of search "8mers_est" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "pappu518.key":
8mers (NA) ID 8mers NA preliminary pattern
and-or
followed by
g or a
g or a
cg
t or c
t or c
c
g or c
followed by
g or c
g or c
g
g or a
g or a
cg
t or c
t or c

Selected files:
File : simplex_est.seq
-- Output Parameters --
Format Options:
Nucleic acid code matching Exact Indirect file NO
Find non-matching hits only No Sequence or key file NO
Report key used Yes List of hits Yes
Note position of hit Yes Hit display Yes
Display full annotations Yes Name and annotations Yes
Sequence context 50

-- Run Parameters --

-- Run Parameters --

Run mode Batch
Time to start comparison now
Notify at end of run No

1 match found in sequence:
bf169293 ; TOIG of: bf169293 check: 1397 from: 1 to: 592
(from "simplex_est.seq")
TOIG of: bf169293 check: 1397 from: 1 to: 592

LOCUS BF169293 592 bp mRNA EST 30-OCT-2000
DEFINITION Tm_ad_04B07_SKPL Trichuris muris (parasitic nematode) mixed adult
Trichuris muris cDNA clone Tm_ad_04B07 5' similar to
emb|CAB93501.1| (Y19221) tropomyosin - Anisakis simplex, mRNA
sequence.
ACCESSION BF169293
VERSION BF169293.1 GI:11053904
KEYWORDS EST.
SOURCE Trichuris muris.
ORGANISM Trichuris muris.
Eukaryota; Metazoa; Nematoda; Trichocephalida; Trichuridae

REFERENCE 1 (bases 1 to 592)
AUTHORS Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guillian,D., Hall
,N., Quayle,M. and Barrell,B.

1 (bases 1 to 497)
AUTHORS Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guillian,D., Hall
,N., Quayle,M. and Barrell,B.

TITLE Edinburgh University/Sanger Centre Nematode EST Project
JOURNAL Unpublished (2000)
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The library was prepared by Richard Grencis, Manchester University,
Manchester. Sequencing was performed by Claire Whitton, ICAPP,
Edinburgh
PCR Primers
FORWARD: T3
BACKWARD: T7PL
Plate: 04 row: B column: 07
Seq primer: SKPL
High quality sequence stop: 311.
Location/Qualifiers
1..592
/organism="Trichuris muris"
/db_xref="taxon:70415"
/clone_lib="Tm_ad_04B07"
/clone_lib="Trichuris muris (parasitic nematode) mixed
adult"
/sex="mixed"
/dev_stage="adult"
/note="Vector: Lambda Zap II; Site_1: EcoRI (5'end);
Site_2: XhoI (3'end); Trichuris muris is a nematode
parasite of rodents related to the human whipworm
Trichuris trichiura. The library was constructed from
Trichuris muris adults (Edinburgh 'E' strain) maintained
in mice, and was provided by Dr. Richard Grencis,
University of Manchester."

BASE COUNT 163 a 141 c 150 g 138 t
ORIGIN
BF169293 Length: 592 March 12, 2002 09:13 Type: N Check: 1397
Found using '8mers' (pappu518.key)

...
167 CTGCTCCAGCAGGAGAGAGTACAGGCCATATCCGAGGAGCTGGATCAGACGTTCCAA
217 224
|-----|
227 GAATTGCTCGCTATTATGAGTCTCCAGTTGGAGTAGAGGGCGCTCTT
1 match found in sequence:
bg577835 ; TOIG of: bg577835 check: 8229 from: 1 to: 497
(from "simplex_est.seq")
TOIG of: bg577835 check: 8229 from: 1 to: 497

LOCUS BG577835 497 bp mRNA EST 11-APR-2001
DEFINITION Tm_ad_12C11_SKPL Trichuris muris (parasitic nematode) mixed adult
Trichuris muris cDNA clone Tm_ad_12C11 5' similar to
emb|CAB93501.1| (Y19221) tropomyosin - Anisakis simplex, mRNA
sequence.
ACCESSION BG577835
VERSION BG577835.1 GI:13592899
KEYWORDS EST.
SOURCE Trichuris muris.
ORGANISM Trichuris muris.
Eukaryota; Metazoa; Nematoda; Trichocephalida; Trichuridae

REFERENCE 1 (bases 1 to 497)
AUTHORS Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guillian,D., Hall
,N., Quayle,M. and Barrell,B.

TITLE Edinburgh University/Sanger Centre Nematode EST Project

JOURNAL
COMMENT

Unpublished (2000)
Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The library was prepared by Richard Grencis, Manchester University,
Manchester. Sequencing was performed by the Pathogen Sequencing
Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart
Barrell).

PCR Primers

FORWARD: T3
BACKWARD: T7PL
Plate: 12 row: C column: 11

Seq primer: SKPL

High quality sequence stop: 486.

FEATURES

source

1. .497
/organism="Trichuris muris"
/db_xref="taxon:70415"
/clone_lib="Tm_ad_12C11"
adult
/sex="mixed"
/dev_stage="adult"
/note="Vector: Lambda Zap II; Site_1: EcoRI (5'end);
Site_2: XhoI (3'end); Trichuris muris is a nematode
parasite of rodents related to the human whipworm
Trichuris trichiura. The library was constructed from
Trichuris muris adults (Edinburgh 'E' strain) maintained
in mice, and was provided by Dr. Richard Grencis,
University of Manchester."
123 a 131 c 122 g 121 t

BASE COUNT
ORIGIN

BG577835 Length: 497 March 12, 2002 09:13 Type: N Check: 8229 ..
Found using '8mers' (pappu518.key)

...

102 CTGCTCCACGAGAGAGATACAGGCCATATCCGAGGAGCTGGATCAGAGCTTCCAA
152 159
|-----|

162 GAATTGCTCGCTATTATAGTGTCCAGTTGGAGTAGAGGGCGTCTTT

...

-- Search Statistics --

Times: CPU Total Elapsed
00:00:00.00 00:00:00.00

Number of sequences searched: 24
Number of sequence hits: 2
Number of separate matches: 2
Number of sequence hits saved: 0

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 19:41:02 ; Search time 2230.08 seconds
(without alignments)
106.008 Million cell updates/sec

Title: US-09-802-518-1

Perfect score: 22

Sequence: 1 tgactgtgaacgttcgagatga 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estpl:*
5: em_estba:*
6: em_estro:*
7: em_estov:*
8: em_estbc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	17.8	80.9	561	13	AZ755668
C 2	17.8	80.9	961	11	BF971856
C 3	17.4	79.1	489	13	AZ060178
C 4	17.4	79.1	530	13	AZ060178
5	17.2	78.2	374	13	AZ245026
6	17.2	78.2	408	13	AZ236502
7	17.2	78.2	424	10	BE723539
8	17.2	78.2	463	10	AU083559
9	17.2	78.2	479	10	AU083559
C 10	17.2	78.2	972	13	CNS05PD9
11	16.8	76.4	105	10	AA094019
12	16.8	76.4	523	13	AZ483488

C 13	16.8	76.4	526	13	AZ501799
C 14	16.8	76.4	681	10	AV732648
C 15	16.8	76.4	705	10	AW916461
C 16	16.4	74.5	400	10	AW398307
C 17	16.4	74.5	496	10	AW034934
C 18	16.4	74.5	546	13	AZ058706
C 19	16.4	74.5	554	13	AZ280611
C 20	16.4	74.5	599	13	AZ068022
C 21	16.4	74.5	600	13	AZ976014
C 22	16.4	74.5	633	11	BG570577
C 23	16.4	74.5	747	11	BG127461
C 24	16.4	74.5	934	13	CNS0275C
C 25	16.2	73.6	202	10	AA236074
C 26	16.2	73.6	236	10	BB183285
C 27	16.2	73.6	236	11	BF932252
C 28	16.2	73.6	251	10	BB565758
C 29	16.2	73.6	286	10	AI099019
C 30	16.2	73.6	297	10	AA45764
C 31	16.2	73.6	300	11	CI11370
C 32	16.2	73.6	304	13	AQ581066
C 33	16.2	73.6	317	10	BE119339
C 34	16.2	73.6	360	10	AV189436
C 35	16.2	73.6	398	10	BE428418
C 36	16.2	73.6	401	10	BE275964
C 37	16.2	73.6	416	13	AQ646593
C 38	16.2	73.6	432	11	BG813145
C 39	16.2	73.6	434	11	BF851024
C 40	16.2	73.6	441	10	AI597068
C 41	16.2	73.6	448	10	AW390277
C 42	16.2	73.6	452	11	BG814270
C 43	16.2	73.6	457	10	BE367478
C 44	16.2	73.6	472	13	AQ904593
C 45	16.2	73.6	473	11	R83554

ALIGNMENTS

RESULT 1
AZ755668/c
LOCUS ev02g09.x1 PAX3 CASTING Library 'ev' Homo sapiens genomic clone
DEFINITION ev02g09 random, DNA sequence.
ACCESSION AZ755668
VERSION AZ755668.1 GI:13175090
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 561)
AUTHORS Barber, T.D., Barber, M.C., Tomescu, O., Barr, F., Ruben, S. and Friedman, T.B.
TITLE Cyclic amplification and selection of target genes regulated by PAX3 and PAX3/FKHR in embryogenesis and alveolar rhabdomyosarcoma Unpublished (2000)
JOURNAL Contact: Friedman TB
COMMENT Laboratory of Molecular Genetics
National Institute on Deafness and Other Communication Disorders,
National Institutes of Health
5 Research Court, Room 2A-15, Rockville, MD 20850, USA
Tel: 301 402 7580
Fax: 301 496 7882
Email: friedman@nidcd.nih.gov
Plate: 02 row: g column: 09
Seq primer: -21M13 forward primer (ABI)
Class: random plasmid subclone.
FEATURES
Location/Qualifiers
1..561
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ev02g09"
/clone_lib="PAX3 CASTING Library 'ev'."

AZ501799 IM0340J17
AV732648 AV732648
AW916461 EST347765
AW398307 EST298154
AW034934 EST279163
AZ058706 RPCI-23-4
AZ280611 RPCI-23-1
AZ068022 RPCI-23-4
AZ976014 2M0251J16
BG570577 602591261
BG127461 EST473107
AL115061 Tetraodon
AA236074 zso5a04.r
BB183285 BB183285
BF932252 IL2-WT019
BB565758 BB565758
AI099019 uc02H08.r
AA45764 vc63b04.s
CI11370 C11370 Yuj1
AQ581066 RPCI-11-4
BE119339 UI-R-CA0-
AV189436 AV189436
BE428418 MTDO06.G1
BE275964 601120704
AQ646593 RPCI93-Dp
BG813145 daf71e08
BF851024 IL5-EN008
AI597068 vj44e09.x
AW390277 CM2-ST018
BG814270 daf65c11
BE367478 P11-B-F04
AQ904593 GSSTC0381
R83554 yql2d03.r1

/sex="Male"
 /lab_host="DH10B"
 /note="vector: pGEM-T Easy; Human genomic DNA was partially digested with Sau3AI, ligated to ds linkers, and enriched for binding to human PAX3dQ+ proteins using a whole genome PCR-based strategy. DNA fragments containing putative PAX3dQ+ binding sites were amplified by PCR and cloned into pGEM-T Easy (Promega). The ligation products were transformed into DH10B electrocompetent cells (Life Technologies)."

BASE COUNT 126 a 149 c 152 g 134 t
 ORIGIN

Query Match 80.9%; Score 17.8; DB 13; Length 561;
 Best Local Similarity 90.5%; Pred. No. 1.8e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagatg 21
 |||||
 Db 461 TGACTGTGAACGTTCAGAGATG 441

RESULT 2
 BF971856 961 bp mRNA 22-JAN-2001
 LOCUS 60224044F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4328890 5',
 DEFINITION mRNA sequence.
 ACCESSION BF971856
 VERSION BF971856.1 GI:12339071
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 961)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LucM189 row: h column: 11
 High quality sequence stop: 555.
 Location/Qualifiers
 1..961
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4328890"
 /clone_lib="NIH_MGC_46"
 /tissue_type="leiomyosarcoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: uterus; Vector: pOT57; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

FEATURES

source
 1..961
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4328890"
 /clone_lib="NIH_MGC_46"
 /tissue_type="leiomyosarcoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: uterus; Vector: pOT57; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 205 a 225 c 240 g 291 t
 ORIGIN

Query Match 80.9%; Score 17.8; DB 11; Length 961;
 Best Local Similarity 90.5%; Pred. No. 2.1e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 gactgtgaacgttcgagatga 22
 |||||
 Db 650 GACTGTGAACGTTCGGGATGA 670

RESULT 3
 LOCUS AZ060178/c 489 bp DNA GSS 30-MAR-2000
 DEFINITION RPCI-23-405E23.TJ RPCI-23 Mus musculus genomic clone RPCI-23-405E23
 , DNA sequence.
 ACCESSION AZ060178
 VERSION AZ060178.1 GI:7351427
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 489)
 Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret
 , B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
 and Fraser, C.M.
 Mouse BAC End Sequences from Library RPCI-23
 Unpublished (1999)
 Other_GSSs: RPCI-23-405E23.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 405 row: E column: 23
 Seq primer: SP6
 Class: BAC ends.
 Location/Qualifiers
 1..489
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-405E23"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

FEATURES

source
 1..489
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-405E23"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 162 a 61 c 67 g 198 t 1 others
 ORIGIN

Query Match 79.1%; Score 17.4; DB 13; Length 489;
 Best Local Similarity 94.7%; Pred. No. 2.7e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgaga 19
 |||||
 Db 170 TGACTGTGAACATTTCGAGA 152

RESULT 4
 AZ886419/c

LOCUS A2886419 530 bp DNA GSS 05-MAR-2001
 DEFINITION RPCI-23-18216.TJ RPCI-23 Mus musculus genomic clone RPCI-23-18216,
 DNA sequence.
 ACCESSION A2886419
 VERSION A2886419.1 GI:13205364
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 530)
 AUTHORS Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret
 B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
 and Fraser, C.M.
 TITLE Mouse BAC End Sequences from Library RPCI-23
 JOURNAL Unpublished (1999)
 COMMENT Other_GSSs: RPCI-23-18216.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC
 library availability, please contact Pieter de Jong
 (pdejong@mail.cho.org). Clones may be purchased from BACPAC
 Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end
 page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
 Plate: 182 row: I column: 6
 Seq primer: SP6
 Class: BAC ends.

FEATURES

source Location/Qualifiers
 1..530
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-18216"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
 EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
 brain genomic DNA was isolated and partially digested
 with a combination of EcoRI and EcoRI Methylase. Size
 selected DNA was cloned into the pBACe3.6 vector at the
 EcoRI sites. The ligation products were transformed into
 DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 174 a 65 c 73 g 218 t
 ORIGIN

Query Match 79.18; Score 17.4; DB 13; Length 530;
 Best Local Similarity 94.78; Pred. No. 2.8e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgaga 19
 |||||
 Db 189 TGACTGTGACATTCGAGA 171

RESULT 5
 A0245026 374 bp DNA GSS 03-OCT-1998
 LOCUS HS_2056_B1_E03_MR CIT Approved Human Genomic Sperm Library D Homo
 DEFINITION sapiens genomic clone Plate=2056 Col=5 Row=J, DNA sequence.
 ACCESSION A0245026
 VERSION A0245026.1 GI:3691600
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Kaller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and
 Hood, L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
 MEDLINE 9380589
 COMMENT

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence tagged connector

Plate: 2056 row: J column: 5

Class: BAC ends

High quality sequence stop: 374.

Location/Qualifiers

source

1..374

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Plate=2056 Col=5 Row=J"

/clone_lib="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in

E-Coli DH10B"

BASE COUNT 91 a 73 c 92 g 117 t 1 others

ORIGIN

Query Match 78.28; Score 17.2; DB 13; Length 374;
 Best Local Similarity 86.48; Pred. No. 3.2e+02;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagatga 22
 |||||

Db 207 TGACTGTGACGATTCGATCA 228

RESULT

AZ536502 408 bp DNA GSS 03-NOV-2000
 LOCUS 110300.96 Planococcus lilacinus DNA Planococcus lilacinus genomic,
 DEFINITION DNA sequence.
 ACCESSION AZ536502
 VERSION AZ536502.1 GI:11093449
 KEYWORDS GSS.
 SOURCE lilac mealybug.
 ORGANISM Planococcus lilacinus

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha;
 Aphidiformes; Coccoidea; Pseudococcidae; Planococcus.

1 (bases 1 to 408)

Mohan, K.N. and Chandra, H.S.

Mealybug shotgun sequencing

Unpublished (2000)

Contact: Mohan KN

Microbiology and Cell Biology

Indian Institute of Science

Sir C.V. Raman Avenue, Bangalore, Karnataka 560012, India

Email: mohan@cmbi.iisc.ernet.in

Class: shotgun.

Location/Qualifiers

source

1..408

/organism="Planococcus lilacinus"

/db_xref="taxon:40930"

/clone_lib="Planococcus lilacinus DNA"

BASE COUNT 134 a 83 c 80 g 111 t

ORIGIN

```

Query Match      78.2%; Score 17.2; DB 13; Length 408;
Best Local Similarity 86.4%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
|||||
Db 251 TGACTGTGAACGACGATATGA 272

RESULT 7
BE723539
LOCUS BE723539 424 bp mRNA EST 25-APR-2001
DEFINITION 193384 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE723539
VERSION BE723539.1 GI:10124826
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 424)
Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett
, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-Mckown, C.G.,
Pertea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
Keefe, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 92 row: E column: 14
Seq primer: ATTTAGGTGACACTATAG.
Location/Qualifiers
1. 424
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 4BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector; pCMV SPORT6; Site.1: XbaI; Site.2: XhoI;
Library made from pooled tissue from day 20 and day 40
embryos."
BASE COUNT 107 a 107 c 87 g 123 t
ORIGIN

Query Match      78.2%; Score 17.2; DB 10; Length 424;
Best Local Similarity 86.4%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
|||||
Db 268 TGAGTCTGACGTTAGACATGA 289

RESULT 8
AU083559
LOCUS AU083559 463 bp mRNA EST 21-MAR-2000
DEFINITION AU083559 Rice green shoot Oryza sativa cDNA clone SI4862, mRNA
sequence.

```

```

ACCESSION AU083559
VERSION AU083559.1 GI:7274015
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 463)
Sasaki, T. and Yamamoto, K.
Rice cDNA from green shoot (2000)
Unpublished (2000)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT "RGP".
SI4862_62.
Location/Qualifiers
1. 463
/organism="Oryza sativa"
/strain="Nipponbare"
/db_xref="taxon:4530"
/clone="SI4862"
/clone_lib="Rice green shoot"
/note="Green shoot (8 days old)"
BASE COUNT 151 a 74 c 109 g 127 t 2 others
ORIGIN

Query Match      78.2%; Score 17.2; DB 10; Length 463;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
|||||
Db 151 TGAGTGTCAATGTTAGACATGA 172

RESULT 9
AU089685
LOCUS AU089685 479 bp mRNA EST 27-APR-2000
DEFINITION AU089685 Rice callus Oryza sativa cDNA clone C40060, mRNA sequence.
ACCESSION AU089685
VERSION AU089685.1 GI:7652165
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 479)
Sasaki, T. and Yamamoto, K.
Rice cDNA from callus (2000)
Unpublished (2000)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT "RGP".
C40060_32.
Location/Qualifiers
1. 479
/organism="Oryza sativa"
/strain="cultivar Nipponbare, sub_species Japonica"
/db_xref="taxon:4530"
/clone="C40060"

```

```

/clone_lib="Rice callus"
/notes="vector: pBluescript II SK+; Site_1: SalI; Site_2:
NotI; cDNA prepared from rice callus mRNAs by using
oligo(dr) as a primer and ligating to the SalI-NotI site
of pBluescript II SK+ phagemid."
BASE COUNT      149 a      96 c      104 g      130 t
ORIGIN

Query Match      78.2%; Score 17.2; DB 10; Length 479;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagatga 22
||||| ||||| ||||| |||||
Db 299 TGAGTGTGAAGTTCAGATGA 320

RESULT 10
LOCUS      CNS05PD9      972 bp      DNA      GSS      26-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
005F08 of library A from tetraodon nigroviridis, genomic survey
sequence.
ACCESSION  AL347814.1 GI:8241584
VERSION     GSS: genome survey sequence.
KEYWORDS    Tetraodon nigroviridis.
SOURCE      Tetraodon nigroviridis
ORGANISM    Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1 (bases 1 to 972)
AUTHORS     Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billaut,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
TITLE       Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
JOURNAL     Unpublished
AUTHORS     2 (bases 1 to 972)
            Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
TITLE       Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
JOURNAL     Unpublished
AUTHORS     3 (bases 1 to 972)
            Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/tetraodon.
FEATURES
            Location/Qualifiers
            source
            1. 972
                /organism="Tetraodon nigroviridis"
                /db_xref="taxon:99883"
                /clone="005F08"
                /clone_lib="A"
                /note="Genoscope sequence ID : COAA005DC04C1-end : T7"
BASE COUNT      195 a      268 c      199 g      297 t      13 others
ORIGIN

Query Match      78.2%; Score 17.2; DB 13; Length 972;
Best Local Similarity 86.4%; Pred. No. 4.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagatga 22
||||| ||||| ||||| |||||
Db 46 TGGCTGTGAAGTTCGAGATGA 25

```

```

RESULT 11
LOCUS      AA094019      105 bp      mRNA      EST      25-OCT-1996
DEFINITION c11619, seq.F Human fetal heart, Lambda ZAP Express Homo sapiens
CDNA 5', mRNA sequence.
ACCESSION  AA094019
VERSION     AA094019.1 GI:1639612
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 105)
AUTHORS     Liew,C.C.
TITLE       CDNAS from fetal heart (1996)
JOURNAL     Unpublished (1996)
COMMENT     Contact: Liew CC
            Department of Laboratory Medicine and Pathobiology
            University of Toronto
            Banning Institute, 100 College St., Toronto, Ontario, M5G1L5
            Tel: 4169788758
            Fax: 4169785650
            Email: liewc@utcc.utoronto.ca
            PCR Primers
            FORWARD: 5' GCCAAGCTCGAATTAACCTCACTAAAGG 3'
            BACKWARD: 5' CCAGTGAATGTGAATACGACTCACTATAGGCG 3'
            Seq primer: 5' GAAATTAACCTCACTAAAGG 3'
            Location/Qualifiers
            1. 105
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="Human fetal heart, Lambda ZAP Express"
                /lab_host="E. coli XL1-Blue"
                /note="vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
                XhoI; mRNA was purified from human fetal hearts (8-10
                weeks). CDNA was synthesized using a XhoI-Oligo dT
                adaptor-primer. EcoRI adaptors were ligated, followed by
                digestion with XhoI, for directional cloning into
                predigested lambda ZAP Express."
BASE COUNT      31 a      21 c      16 g      37 t
ORIGIN

Query Match      76.4%; Score 16.8; DB 10; Length 105;
Best Local Similarity 90.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tgactgtgaacgttcgagat 20
||||| ||||| ||||| |||||
Db 43 TGACTGTGAACCTTCAGAT 62

RESULT 12
LOCUS      AZ483488      523 bp      DNA      GSS      05-OCT-2000
DEFINITION 1M0309M12F Mouse 10kb plasmid UUC1M library Mus musculus genomic
clone UUC1M0309M12 F, DNA sequence.
ACCESSION  AZ483488
VERSION     AZ483488.1 GI:10647510
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 523)
AUTHORS     Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            M., Rose,M., Rose,R., Stokes,R., Tingey,A., von-Niederhausern,A.
            and Wright,D., Weiss,R.
TITLE       Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts

```

JOURNAL
COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0309 row: M column: 12
Seq primer: CTTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 523.

FEATURES

source

1. 523
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC1M0309M12"
/clone_lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 162 a 127 c 103 g 131 t
ORIGIN

Query Match 76.4%; Score 16.8; DB 13; Length 523;
Best Local Similarity 90.0%; Pred. No. 5.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 actgtgaacgttcgagatga 22
||||| || ||||| |||||
Db 376 ACTGTGTACTTCGAGATGA 395

RESULT 13

AZ501799 526 bp DNA GSS 05-OCT-2000
LOCUS 1M0340J17R Mouse 10kb plasmid UUC1M library Mus musculus genomic
DEFINITION clone UUC1M0340J17 R, DNA sequence.
ACCESSION AZ501799
VERSION AZ501799.1 GI:10683115
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 1 to 526)
Dunn, B., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb

TITLE

JOURNAL
COMMENT

plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0340 row: J column: 17
Seq primer: CACACGAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 526.

FEATURES

source

1. 526
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC1M0340J17"
/clone_lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 121 a 126 c 139 g 140 t
ORIGIN

Query Match 76.4%; Score 16.8; DB 13; Length 526;
Best Local Similarity 90.0%; Pred. No. 5.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgactgtgaacgttcgagat 20
||||| ||||| ||||| |||||
Db 60 TGACAGTGAACGTTCTAGAT 79

RESULT 14

AV732648/c 681 bp mRNA EST 17-OCT-2000
LOCUS AV732648 HTF Homo sapiens cDNA clone HTFBLB03 5', mRNA sequence.
DEFINITION AV732648
ACCESSION AV732648
VERSION AV732648.1 GI:10850193
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 681)
Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Ou, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z., and Han, Z.
Homo sapiens cDNA HTF clones

TITLE

JOURNAL
COMMENT

Unpublished (2000)
Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

source

1. .681
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HTFBLB03"
/clone_lib="HTF"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 202 a 137 c 156 g 184 t 2 others
ORIGIN

Query Match

Best Local Similarity 76.4%; Score 16.8; DB 10; Length 681;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 actgtgaacgttcgagatga 22

||||| ||| |||||

Db 382 ACTCTGAACATTTGAGATGA 363

RESULT 15

AW916461

LOCUS

AW916461 705 bp mRNA EST 25-MAY-2000
EST347765 Rat gene index, normalized rat, norvegicus, Bento Soares
Rattus norvegicus cDNA clone RGIDQ49 5' end, mRNA sequence.

ACCESSION

AW916461

VERSION

AW916461.1 GI:8082187

KEYWORDS

EST.

SOURCE

Norway rat.

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 705)

AUTHORS

Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavage, A.R. and Adams, M.D.

TITLE

Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat

JOURNAL

Gene Index

COMMENT

Unpublished (1998)

Contact: Lee, NH

The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301)-838-3529

Fax: (301)-838-0208

Email: nhlee@tigr.org

This clone is available through the ATCC, contact the ATCC

tel#703-365-2700 for further information

Seq primer: M13 Reverse.

FEATURES

source

1. .705
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="RGIDQ49"
/clone_lib="Rat gene index, normalized rat, norvegicus,
Bento Soares"
/tissue_type="mix - brain, ovary, placenta, kidney, lung,
liver, embryo, heart, muscle, spleen"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Estimated insert size approx.1 kb"

BASE COUNT
ORIGIN

177 a 201 c 169 g 157 t 1 others

Query Match 76.4%; Score 16.8; DB 10; Length 705;
Best Local Similarity 90.0%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 actgtgaacgttcgagatga 22

||||| ||| |||||

Db 150 ACTGTGACCTTCGAGATGA 169

Search completed: March 11, 2002, 20:30:00
Job time: 2938 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 19:51:12 ; Search time 100.2 Seconds
(without alignments)
49.726 Million cell updates/sec

Title: US-09-802-518-1

Perfect score: 22
Sequence: 1 tgactgtgaacgttcgagatga 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCRUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20.4	92.7	22	4	US-09-092-314-2
2	18.8	85.5	22	4	US-09-092-314-1
3	18.8	85.5	22	4	US-09-092-314-3
4	18.8	85.5	22	4	US-09-092-314-10
5	17.2	78.2	22	4	US-09-092-314-4
6	15.6	70.9	22	4	US-09-092-314-5
7	15.6	70.9	22	4	US-09-092-314-7
8	15.6	70.9	22	4	US-09-092-314-8
9	15.6	70.9	1418	1	US-08-391-615-7
10	15.6	70.9	1830	1	US-09-019-931-2
11	15.6	70.9	2505	1	US-08-391-615-1
12	15.6	70.9	6909	2	US-08-804-196-1
13	15.6	70.9	6909	2	US-08-658-340-1
14	15.6	70.9	6909	3	US-08-746-111-26
15	15.2	69.1	1892	2	US-08-933-750C-66
16	15.2	69.1	1892	3	US-09-234-613-66
17	15.2	69.1	6638	2	US-08-070-301-2
18	14.8	67.3	882	1	US-08-622-354-4
19	14.6	66.4	404	4	US-09-060-756-303
20	14.6	66.4	913	2	US-08-975-316-61
21	14.6	66.4	1532	3	US-09-118-324-1
22	14.6	66.4	1614	4	US-09-046-894-29
23	14.6	66.4	2694	3	US-08-975-703-5
24	14.6	66.4	2694	4	US-09-515-884-5
25	14.6	66.4	5408	1	US-08-471-058-20
26	14.6	66.4	5408	3	US-08-471-057-20
27	14.6	66.4	4403765	4	US-09-103-840A-2

28	14.6	66.4	4411529	4	US-09-103-840A-1
c 29	14.2	64.5	864	4	US-08-998-416-297
30	14.2	64.5	1166	1	US-08-121-063-7
c 31	14.2	64.5	176373	3	US-09-128-155-17
32	14	63.6	77	1	US-08-399-412A-58
33	14	63.6	95	4	US-08-952-793-258
34	14	63.6	95	5	PCT-US96-09455A-258
35	14	63.6	657	2	US-08-479-733A-22
36	14	63.6	657	3	US-08-487-427-22
37	14	63.6	657	3	US-08-479-727A-22
38	14	63.6	657	3	US-08-482-369A-22
39	14	63.6	657	3	US-09-009-217-10
40	14	63.6	657	3	US-09-009-656-10
41	14	63.6	657	5	PCT-US95-07439-22
42	14	63.6	672	1	US-07-816-679A-2
43	14	63.6	672	5	PCT-US92-11270-2
c 44	14	63.6	753	4	US-08-861-774E-73
45	14	63.6	789	4	US-08-776-059-42

ALIGNMENTS

RESULT 1
US-09-092-314-2
; Sequence 2, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; TITLE OF INVENTION: Sequence Activity
; Patent No. 6225292
; FILE REFERENCE: 6510-173US1
; CURRENT APPLICATION NUMBER: US/09/092,314
; CURRENT FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,794
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-092-314-2

Query Match 92.7%; Score 20.4; DB 4; Length 22;
Best Local Similarity 95.5%; Pred. No. 0.049;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 tgactgtgaacgttcgagatga 22
RESULT 2
US-09-092-314-1
; Sequence 1, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; TITLE OF INVENTION: Sequence Activity
; Patent No. 6225292
; FILE REFERENCE: 6510-173US1
; CURRENT APPLICATION NUMBER: US/09/092,314
; CURRENT FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,794
; PRIOR FILING DATE: 1997-06-06

```
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-092-314-1
```

```
Query Match      85.5%; Score 18.8; DB 4; Length 22;
Best Local Similarity 90.9%; Pred. No. 0.33;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 tgactgtgaacgttcgagatga 22
   ||||| ||||| ||||| |||||
Db 1 tgactgtgaaggttagagatga 22
```

```
RESULT 3
US-09-092-314-3
; Sequence 3, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; APPLICANT: Roman, Mark
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; TITLE OF INVENTION: Sequence Activity
; Patent No. 6225292
; FILE REFERENCE: 6510-173US1
; CURRENT APPLICATION NUMBER: US/09/092.314
; CURRENT FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,794
; PRIOR FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-092-314-3
```

```
Query Match      85.5%; Score 18.8; DB 4; Length 22;
Best Local Similarity 90.9%; Pred. No. 0.33;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 tgactgtgaacgttcgagatga 22
   ||||| ||||| ||||| |||||
Db 1 tgactgtgaaccttagagatga 22
```

```
RESULT 4
US-09-092-314-10
; Sequence 10, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; APPLICANT: Roman, Mark
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; TITLE OF INVENTION: Sequence Activity
; Patent No. 6225292
; FILE REFERENCE: 6510-173US1
; CURRENT APPLICATION NUMBER: US/09/092.314
; CURRENT FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,794
; PRIOR FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
```

```
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-092-314-10
```

```
Query Match      85.5%; Score 18.8; DB 4; Length 22;
Best Local Similarity 90.9%; Pred. No. 0.33;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 tgactgtgaacgttcgagatga 22
   ||||| ||||| ||||| |||||
Db 1 tgactgtgaatgttagagatga 22
```

```
RESULT 5
US-09-092-314-4
; Sequence 4, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; APPLICANT: Roman, Mark
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; TITLE OF INVENTION: Sequence Activity
; Patent No. 6225292
; FILE REFERENCE: 6510-173US1
; CURRENT APPLICATION NUMBER: US/09/092.314
; CURRENT FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,794
; PRIOR FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-092-314-4
```

```
Query Match      78.2%; Score 17.2; DB 4; Length 22;
Best Local Similarity 86.4%; Pred. No. 2.2;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 tgactgtgaacgttcgagatga 22
   ||||| ||||| ||||| |||||
Db 1 tgactgtgaagcttagagatga 22
```

```
RESULT 6
US-09-092-314-5
; Sequence 5, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; APPLICANT: Roman, Mark
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; TITLE OF INVENTION: Sequence Activity
; Patent No. 6225292
; FILE REFERENCE: 6510-173US1
; CURRENT APPLICATION NUMBER: US/09/092.314
; CURRENT FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,794
; PRIOR FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-092-314-5

Query Match 70.9%; Score 15.6; DB 4; Length 22;
Best Local Similarity 81.8%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
||||||| | | | | | | |
Db 1 tgactgtgtctcttagagatga 22

RESULT 7
US-09-092-314-7
; Sequence 7, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; APPLICANT: Roman, Mark
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; TITLE OF INVENTION: Sequence Activity
; Patent No. 6225292
; FILE REFERENCE: 6510-173US1
; CURRENT APPLICATION NUMBER: US/09/092.314
; CURRENT FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,794
; PRIOR FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-092-314-7

Query Match 70.9%; Score 15.6; DB 4; Length 22;
Best Local Similarity 81.8%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
||||||| | | | | | | |
Db 1 tgactgtgaggtcagagatga 22

RESULT 8
US-09-092-314-8
; Sequence 8, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; APPLICANT: Roman, Mark
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; TITLE OF INVENTION: Sequence Activity
; Patent No. 6225292
; FILE REFERENCE: 6510-173US1
; CURRENT APPLICATION NUMBER: US/09/092.314
; CURRENT FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,794
; PRIOR FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-092-314-8

Query Match 70.9%; Score 15.6; DB 4; Length 22;
Best Local Similarity 81.8%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 tgactgtgaacgttcgagatga 22
||||||| | | | | | | |
Db 1 tgactgtgagggctagagatga 22

RESULT 9
US-08-391-615-7
; Sequence 7, Application US/08391615
; Patent No. 5550054
; GENERAL INFORMATION:
; APPLICANT: Witte, Owen
; APPLICANT: Tsukada, Satoshi
; APPLICANT: Saffran, Douglas
; APPLICANT: Rawlings, David
; TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/391.615
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/006,449
; FILING DATE: 21-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-57583-1/BIR UCLA 232-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1418 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-391-615-7

Query Match 70.9%; Score 15.6; DB 1; Length 1418;
Best Local Similarity 81.8%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 tgactgtgaacgttcgagatga 22
||||||| | | | | | | |
Db 218 TGACTTTGACCTGGGAGAAGA 239

RESULT 10
US-09-019-931-2
; Sequence 2, Application US/09019931
; Patent No. 6194148

;; GENERAL INFORMATION:
;; APPLICANT: Hori, Kunio, Takahashi, Takeo, Okada, Takao
;; TITLE OF INVENTION: A Method For Detecting A Hybridized
;; TITLE OF INVENTION: Nucleic Acid Molecule
;; NUMBER OF SEQUENCES: 2
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Frishauf, Holtz, Goodman, Langer & Chick
;; STREET: 767 Third Avenue
;; CITY: New York
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10017-2023
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3+ inch, 1.44 mb diskette
;; COMPUTER: IBM PC
;; OPERATING SYSTEM: MS DOS
;; SOFTWARE: ASCII
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/019,931
;; FILING DATE: 06-FEB-1998
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 9-025291
;; FILING DATE: 07-FEB-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Barth, Richard
;; REGISTRATION NUMBER: 28,180
;; REFERENCE/DOCKET NUMBER: 980048/HG
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 319-4900
;; TELEFAX: (212) 319-3101
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1830
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: other nucleic acid
;; IMMEDIATE SOURCE:
;; LIBRARY: pGEX-PH
;; US-09-019-931-2

Query Match 70.9%; Score 15.6; DB 4; Length 1830;
Best Local Similarity 81.8%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
||||| ||||| ||||| ||
Db 881 TGACTTTGACGTGGGAGA 902

RESULT 11
US-08-391-615-1
; Sequence 1, Application US/08391615
; Patent No. 5550054
; GENERAL INFORMATION:
; APPLICANT: Witte, Owen
; APPLICANT: Tsukada, Satoshi
; APPLICANT: Saffran, Douglas
; APPLICANT: Rawlings, David
; TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/391,615
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/006,449
;; FILING DATE: 21-JAN-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Rowland, Bertram I
;; REGISTRATION NUMBER: 20,015
;; REFERENCE/DOCKET NUMBER: A-57583-1/BIR UCLA 232-1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 781-1889
;; TELEFAX: (415) 398-3249
;; TELEX: 910 277299 FHT UR
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2505 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: CDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 137..2116
;; US-08-391-615-1

Query Match 70.9%; Score 15.6; DB 1; Length 2505;
Best Local Similarity 81.8%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
||||| ||||| ||||| ||
Db 262 TGACTTTGACGTGGGAGA 283

RESULT 12
US-08-804-196-1/c
; Sequence 1, Application US/08804196
; Patent No. 5874256
; GENERAL INFORMATION:
; APPLICANT: Bertina, Rogier
; APPLICANT: Reitsma, Pieter
; TITLE OF INVENTION: A method for diagnosing an increased
; TITLE OF INVENTION: risk for thrombosis or a genetic defect causing
; TITLE OF INVENTION: thrombosis and kit for use with the same.
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5874256el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,196
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/454,353
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Gortley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 948-7400
TELEFAX: (301) 948-9751
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6909 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
US-08-804-196-1

Query Match 70.9%; Score 15.6; DB 2; Length 6909;
Best Local Similarity 81.8%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
||||||| ||| |||||

Db 793 TGACTGTGATACATTAGGATGA 772

RESULT 13

US-08-658-340-1/c
; Sequence 1, Application US/08658340
; Patent No. 5910576
; Patent No. 5910576 5861489
; GENERAL INFORMATION:
; APPLICANT: Bertina, Rogier
; APPLICANT: Reitsma, Pieter
; TITLE OF INVENTION: A method for diagnosing an increased
; TITLE OF INVENTION: risk for thrombosis or a genetic defect causing
; TITLE OF INVENTION: thrombosis and kit for use with the same.
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arzo No. 5910576 5861489el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/658,340
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/454,353
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 948-7400
; TELEFAX: (301) 948-9751
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6909 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
US-08-658-340-1

Query Match 70.9%; Score 15.6; DB 2; Length 6909;
Best Local Similarity 81.8%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
||||||| ||| |||||

Db 793 TGACTGTGATACATTAGGATGA 772

RESULT 14

US-08-746-111-26/c
; Sequence 26, Application US/08746111
; Patent No. 6066778
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Cui, Jisong
; TITLE OF INVENTION: Compositions And Methods For Screening
; TITLE OF INVENTION: Compounds For Anticoagulant Activity
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/746,111
; FILING DATE: 06-NOV-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: UM-02536
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6909 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-746-111-26

Query Match 70.9%; Score 15.6; DB 3; Length 6909;
Best Local Similarity 81.8%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 tgactgtgaacgttcgagatga 22
||||||| ||| |||||

Db 793 TGACTGTGATACATTAGGATGA 772

RESULT 15

US-08-933-750C-66/c
; Sequence 66, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive

```

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750C
FILING DATE: September 23, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 1892 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTNOT03
CLONE: 641127
US-08-933-750C-66

```

```

Query Match      69.1%  Score 15.2; DB 2; Length 1892;
Best Local Similarity 85.0%; Pred. NO. 45;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  tgactgtgaacgttcgagat 20
        ||| ||||| ||||| |
Db      1278  TGAATGTGAAGGTTTCGAGCT 1259

```

Search completed: March 11, 2002, 21:01:45
Job time: 4233 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 20:55:45 ; Search time 1535.11 seconds
(without alignments)
257.918 Million cell updates/sec

Title: us-09-802-518-9

Perfect score: 24

Sequence: 1 tcgtcgaaacgttcgttaacgttcg 24

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_om.*
- 20: em_or.*
- 21: em_ov.*
- 22: em_pat.*
- 23: em_ph.*
- 24: em_pl.*
- 25: em_ro.*
- 26: em_sts.*
- 27: em_sy.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htgo_hum.*
- 31: em_htgo_inv.*
- 32: em_htgo_rod.*
- 33: em_htg_hum.*
- 34: em_htg_inv.*
- 35: em_htg_rod.*
- 36: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	18.4	76.7	148263	8	AC051624	AC051624 Genomic S
C 2	18.2	75.8	1063	1	AB011203	AB011203 Streptococ
C 3	18.2	75.8	166050	2	AC079766	AC079766 Homo sapi
C 4	17.6	73.3	1680	1	HMORFA	X73823 H.mediterra
C 5	17.6	73.3	43075	7	U88974	U88974 Streptococc
C 6	17.6	73.3	169696	2	AC083943	AC083943 Oryza sat
C 7	17.4	72.5	121524	8	AF041468	AF041468 Gullardi
C 8	17.2	71.7	5905	5	AF007218	AF007218 Tetraodon
C 9	17.2	71.7	183317	2	CNS07EF1	AL513003 Oryza sat
C 10	17	70.8	224	8	AY024027	AY024027 Oryza sat
C 11	17	70.8	101803	2	AP003370	AP003370 Oryza sat
C 12	17	70.8	104718	2	AP003987	AP003987 Oryza sat
C 13	17	70.8	149498	2	AP003196	AP003196 Oryza sat
C 14	17	70.8	150379	8	AP003074	AP003074 Oryza sat
C 15	17	70.8	150641	2	AP003211	AP003211 Oryza sat
C 16	16.8	70.0	81580	8	AP000736	AP000736 Arabidops
C 17	16.8	70.0	98017	8	AC027033	AC027033 Arabidops
C 18	16.8	70.0	136047	36	AC069470	AC069470 Arabidops
C 19	16.8	70.0	140019	2	AC093178	AC093178 Oryza sat
C 20	16.8	70.0	347950	1	AP003013	AP003013 Mesorhizo
C 21	16.6	69.2	645	3	SMU35432	U35432 Schistosoma
C 22	16.6	69.2	766	3	ASNOS09	AF130132 Anopheles
C 23	16.6	69.2	3414	3	DMU21123	U21123 Drosophila
C 24	16.6	69.2	9180	1	AE005246	AE005246 Escherich
C 25	16.6	69.2	10294	1	AE001700	AE001700 Thermocog
C 26	16.6	69.2	28224	2	AC014904	AC014904 Drosophill
C 27	16.6	69.2	39708	3	CEF35G12	Z46242 Caenorhabdi
C 28	16.6	69.2	70283	2	AC092870	AC092870 Takifugu
C 29	16.6	69.2	72019	2	AC009914	AC009914 Drosophill
C 30	16.6	69.2	103683	2	AC012387	AC012387 Drosophill
C 31	16.6	69.2	106659	2	AC020331	AC020331 Drosophill
C 32	16.6	69.2	107967	9	AL353701	AL353701 Human DNA
C 33	16.6	69.2	128444	2	AC019924	AC019924 Drosophill
C 34	16.6	69.2	161172	3	AC007839	AC007839 Drosophill
C 35	16.6	69.2	178692	8	AC024594	AC024594 Oryza sat
C 36	16.6	69.2	180495	3	AC009847	AC009847 Drosophill
C 37	16.6	69.2	188633	3	AC007175	AC007175 Drosophill
C 38	16.6	69.2	201400	2	AC067780	AC067780 Homo sapi
C 39	16.6	69.2	229896	14	AF232689	AF232689 Rat cytom
C 40	16.6	69.2	262278	1	AP002552	AP002552 Escherich
C 41	16.6	69.2	287859	3	AE003797	AE003797 Drosophill
C 42	16.6	69.2	296202	3	AE003785	AE003785 Drosophill
C 43	16.6	69.2	302547	3	AE003453	AE003453 Drosophill
C 44	16.4	68.3	2958	6	E04235	E04235 Temperature
C 45	16.4	68.3	2958	6	E04236	E04236 Temperature

ALIGNMENTS

RESULT 1
AC051624/c

LOCUS

DEFINITION

AC051624

VERSION

KEYWORDS

SOURCE

ORGANISM

Oryza sativa

Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS

Du, H., Minx, P., Abbott, A., Doebber, A., de la Bastide, M., Spiegel, L., Nascimento, L., Preston, R., Kirchoff, K., King, L., Vli, M.D., Baker, J., Zutavern, T., Santos, L., Bell, M., Miller, B., Kuit, K., Rodriguez, S., Cunnius, D.M., Ballja, V., Shah, R., A., O'Shaughnessy, A., Palmer, L., Yang, C., Dedhia, N. and McCombe, W.R.

Genomic Sequence for Oryza sativa, Nipponbare strain, Clone

PLN

22-MAY-2001

Genomic Sequence for Oryza sativa, Nipponbare strain, clone

OSJNB0036B06, from Chromosome 10, complete sequence.

AC051624

HTG.

GI:14165313

OSUNB0036B06, from Chromosome 10, complete sequence
Unpublished
2 (bases 1 to 148263)
McCombie,W.R.
Direct Submission
Submitted (15-APR-2000) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA
3 (bases 1 to 148263)
McCombie,W.R.
Direct Submission
Submitted (22-MAR-2001) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA
On May 22, 2001 this sequence version replaced gi:9972290.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

FEATURES
source
Location/Qualifiers
1..148263
/organism="Oryza sativa"
/strain="Nipponbare"
/db_xref="taxon:4530"
/chromosome="10"
/clone="OSUNB0036B06"
/clone_lib="EcoRI"
37007..37090
/note="The sequence contains a dinucleotide (TA) repeat from base position 37007 to 37090 in which the exact length is unknown. Assembly of the database is consistent with PCR."

misc_feature
36837 a 35840 c 36643 g 38943 t

BASE COUNT 36837 a 35840 c 36643 g 38943 t

ORIGIN

Query Match 76.7%; Score 18.4; DB 8; Length 148263;
Best Local Similarity 95.0%; Pred. No. 74;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 cgaacgttcgttaacgttcg 24
|||||
Db 65607 CGATCGTTCGTTAAGTTCG 65588

RESULT 2
LOCUS AB011203 1063 bp DNA BCT 27-JUL-2000
DEFINITION Streptococcus pneumoniae pbp2X gene for penicillin binding protein 2X, partial cds, isolate #13/S04.
ACCESSION AB011203
VERSION AB011203.1 GI:4519225
KEYWORDS pbp2X; penicillin binding protein 2X.
SOURCE Streptococcus pneumoniae (isolate:#13/S04) DNA.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.

REFERENCE
1 (sites)
Asahi,Y., Takeuchi,Y. and Ubukata,K.
Diversity of substitutions within or adjacent to conserved amino acid motifs of penicillin-binding protein 2X in cephalosporin-resistant Streptococcus pneumoniae isolates Antimicrob. Agents Chemother. 43 (5), 1252-1255 (1999)
99240374
REFERENCE 2 (bases 1 to 1063)
Asahi,Y. and Ubukata,K.
Direct Submission
Submitted (13-FEB-1998) to the DDBJ/EMBL/GenBank databases. Kimiko Ubukata, Mei-ji Seika Kaisha, LTD, Pharmaceutical Research Center; 760 Morooka-cho, Kohoku-ku, Yokohama, Kanagawa 222-8567, Japan

OSUNB0036B06, from Chromosome 10, complete sequence
Unpublished
2 (bases 1 to 148263)
McCombie,W.R.
Direct Submission
Submitted (15-APR-2000) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA
3 (bases 1 to 148263)
McCombie,W.R.
Direct Submission
Submitted (22-MAR-2001) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA
On May 22, 2001 this sequence version replaced gi:9972290.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

FEATURES
source
Location/Qualifiers
1..148263
/organism="Oryza sativa"
/strain="Nipponbare"
/db_xref="taxon:4530"
/chromosome="10"
/clone="OSUNB0036B06"
/clone_lib="EcoRI"
37007..37090
/note="The sequence contains a dinucleotide (TA) repeat from base position 37007 to 37090 in which the exact length is unknown. Assembly of the database is consistent with PCR."

misc_feature
36837 a 35840 c 36643 g 38943 t

BASE COUNT 36837 a 35840 c 36643 g 38943 t

ORIGIN

Query Match 76.7%; Score 18.4; DB 8; Length 148263;
Best Local Similarity 95.0%; Pred. No. 74;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 cgaacgttcgttaacgttcg 24
|||||
Db 65607 CGATCGTTCGTTAAGTTCG 65588

RESULT 2
LOCUS AB011203 1063 bp DNA BCT 27-JUL-2000
DEFINITION Streptococcus pneumoniae pbp2X gene for penicillin binding protein 2X, partial cds, isolate #13/S04.
ACCESSION AB011203
VERSION AB011203.1 GI:4519225
KEYWORDS pbp2X; penicillin binding protein 2X.
SOURCE Streptococcus pneumoniae (isolate:#13/S04) DNA.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.

REFERENCE
1 (sites)
Asahi,Y., Takeuchi,Y. and Ubukata,K.
Diversity of substitutions within or adjacent to conserved amino acid motifs of penicillin-binding protein 2X in cephalosporin-resistant Streptococcus pneumoniae isolates Antimicrob. Agents Chemother. 43 (5), 1252-1255 (1999)
99240374
REFERENCE 2 (bases 1 to 1063)
Asahi,Y. and Ubukata,K.
Direct Submission
Submitted (13-FEB-1998) to the DDBJ/EMBL/GenBank databases. Kimiko Ubukata, Mei-ji Seika Kaisha, LTD, Pharmaceutical Research Center; 760 Morooka-cho, Kohoku-ku, Yokohama, Kanagawa 222-8567, Japan

(E-mail:Kimiko_Ubukata@meiji.co.jp, Tel.:81-45-545-3106, Fax:81-45-545-3129)
Location/Qualifiers
1..1063
/organism="Streptococcus pneumoniae"
/isolate="#13/S04"
/db_xref="taxon:1313"
1..1063
/gene="pbp2X"
<1..>1063
/gene="pbp2X"
/note="corresponding to nucleotides 1018-2080 and codon 256-609 in the pbp2X; EMBL Acc#No.X16367"
/codon_start=1
/transl_table=11
/product="penicillin binding protein 2X"
/protein_id="BAA75526.1"
/db_xref="GI:4519226"
/translation="VDGKVVYTLSSPLQSFMEQDAFOEKVKGYMTATLVSAKGTG EILATTQPTFNADTKGDKITKDFVMDILYQSNYEPGSTNKGVMMLAAIDNKTFFGG EYNSSELKADATIRDMVNEGLTSGCTMTFSOGFAHSSNIGITLLQKMGDATWLDY LNRFKGVPTFRGLTDEYTGQLPADNIVNIAMSAGGQISVTQOMLRAFTAIANDGV MLEPKFISALYDNPNDQSVRKSVQKVVGNPVSKAASVTRDHVMVVGDTPTGTYNHS TGRATVNPQGNVAKSGTQAIDKNGGYLTGSTNYIFSVMHMPAENPDFILYVTV QQPEHYSGIQLGEFANPILE"
244..255
/gene="pbp2X"
/note="STMK homology box"
244..246
/gene="pbp2X"
/note="active site serin; unclassified site"
418..426
/gene="pbp2X"
/note="SSN homology box"
874..882
/gene="pbp2X"
/note="KSG homology box"
303 a 212 c 249 g 299 t

BASE COUNT 303 a 212 c 249 g 299 t

ORIGIN

Query Match 75.8%; Score 18.2; DB 1; Length 1063;
Best Local Similarity 87.0%; Pred. No. 89;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 tcgtcgacgttcgttaacgttc 23
|||||
Db 379 TCGTCAACCTCGTTAAGTCC 357

RESULT 3
LOCUS AC079766 166050 bp DNA HTG 08-MAY-2001
DEFINITION Homo sapiens chromosome RPC1-11 clone RP11-189p17, WORKING DRAFT SEQUENCE, 31 unordered pieces.
ACCESSION AC079766
VERSION AC079766.4 GI:13775303
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 166050)
Waterston,R.H.
The sequence of Homo sapiens clone Unpublished
REFERENCE 2 (bases 1 to 166050)
Waterston,R.H.
Direct Submission
Submitted (10-SEP-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
On Apr 24, 2001 this sequence version replaced gi:12718914.

misc_feature 92588..101334
 /note="assembly_name:Contig47"
 misc_feature 101435..110619
 /note="assembly_name:Contig48"
 misc_feature 110720..120293
 /note="assembly_name:Contig49"
 misc_feature 120394..132322
 /note="assembly_name:Contig50"
 misc_feature 132423..146742
 /note="assembly_name:Contig51"
 misc_feature 146843..166050
 /note="assembly_name:Contig52"
 BASE COUNT 46141 a 35144 c 34884 g 46840 t 3041 others
 ORIGIN

Query Match 75.8%; Score 18.2; DB 2; Length 166050;
 Best Local Similarity 87.08; Pred. No. 94;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Oy 2 cgtcgaaacgttcgtaacgttcg 24
 ||| ||||| || ||||| |||||
 Db 47901 CGTGAACGTTGCTGACAGTTTCG 47923

RESULT 4
 HMORFA HMORFA 1680 bp DNA BCT 27-SEP-1993
 LOCUS H.mediterranei DNA sequence for orf 61.1, 61.2 and 61.3.
 DEFINITION X73823
 ACCESSION X73823.1 GI:313654
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Haloferax mediterranei.
 Haloferax mediterranei
 Archaea: Euryarchaeota; Halobacteriales; Halobacteriaceae;
 Haloferax.
 1 (bases 1 to 1680)
 Mojica, F.J., Juez, G. and Rodriguez-Valera, F.
 Transcription at different salinities of Haloferax mediterranei
 sequences adjacent to partially modified PstI sites
 Mol. Microbiol. 9 (3), 613-621 (1993)
 JOURNAL 94018655
 MEDLINE
 REFERENCE 2 (bases 1 to 1680)
 Mojica, F.M.
 Direct Submission
 Submitted (30-JUN-1993) F.M. Mojica, Universidad de Alicante, Dept
 de Genetica y Microbiologia, Campus de San Juan, Apdo. 374, 03080
 Alicante, SPAIN

FEATURES
 source
 1..1680
 /organism="Haloferax mediterranei"
 /strain="R4"
 /db_xref="taxon:2252"
 complement(303..881)
 /note="orf61.1"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA52042.1"
 /db_xref="GI:313656"
 /db_xref="SPTREMBL:O08159"
 /translation="MKRRSLLLSIGASVGTGAVFGSGAFTSIEADRNVLNKVTDNSG
 SQAQVDSQSGVGADAIIVGTDNSESAAEIKFAETNLNRAKTTFTDALEIDNNGGTVK
 DLYVDSSTEGIGDPSNGVLDPRVDGSSIVGDTSGNAIELGADGSTSGGENTDVIE
 VDIVDLDNDVNDGSSSLKDIDSFTVVEAKRN"
 394..918
 /note="orf61.2"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA52041.1"
 /db_xref="GI:313655"
 /db_xref="SPTREMBL:O08160"
 /translation="MSTSMTSVSPPEVLPSAPSSMAFPDVPSPPTIEDSTRKSRTPS
 LIGSSPIPSVLSTYKSTLVPPLLSSISSVNVVFARSURFVSANLITSADSLLSVPT

CDS
 394..918
 /note="orf61.2"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA52041.1"
 /db_xref="GI:313655"
 /db_xref="SPTREMBL:O08160"
 /translation="MSTSMTSVSPPEVLPSAPSSMAFPDVPSPPTIEDSTRKSRTPS
 LIGSSPIPSVLSTYKSTLVPPLLSSISSVNVVFARSURFVSANLITSADSLLSVPT

IASAPTLPWMSKLTCAADPLSLVTLRLTFRSASIEVNAPEPNTAPVPTIAAPILSRRL
 RFIFLDVSRGTSY"
 993..1643
 /note="orf61.3; differently transcribed depending on the
 salinity of the growth medium"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA52043.1"
 /db_xref="GI:313657"
 /db_xref="SPTREMBL:O08161"
 /translation="MSKMNRRSVLGLIGTIGTGAAGCGAFSTVEATREVEVNVV
 GVDGDDDDGSEIADQFTDLVDASADEVALRDSNDLVDTPTTLPTNDLQANG
 TADIDENYSLIANDVTVEFGYKGGGTDRLPLNPTVYPSGFTLVNNGDNGNGKK
 FSVELESNTFLTEIGQNPHSEVAADTTSERFATLETQCGSESTLTIRISEV"
 BASE COUNT 337 a 478 c 468 g 397 t
 ORIGIN

Query Match 73.3%; Score 17.6; DB 1; Length 1680;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Oy 1 tcgtcgaaacgttcgtaacgttcg 24
 ||| ||||| ||||| |||||
 Db 626 TCTTCGACGTTGCTGAGGTTTCG 649

RESULT 5
 U88974/c
 LOCUS Streptococcus thermophilus Temperate bacteriophage O1205, complete
 DEFINITION genome.
 U88974
 U88974.1 GI:2444080
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Streptococcus thermophilus temperate bacteriophage O1205.
 Streptococcus thermophilus temperate bacteriophage O1205
 Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
 1 (bases 1 to 43075)
 Stanley, E., Fitzgerald, G.F., Le Marrec, C., Fayard, B. and van
 Sinderen, D.
 Sequence analysis and characterization of phi O1205, a temperate
 bacteriophage infecting Streptococcus thermophilus CNR21205
 Microbiology 143 (Pt 11), 3417-3429 (1997)
 JOURNAL 98048466
 MEDLINE
 REFERENCE 2 (bases 1 to 43075)
 Stanley, E., Fitzgerald, G.F., Le Marrec, C., Fayard, B. and van
 Sinderen, D.
 Direct Submission
 Submitted (10-FEB-1997) Department of Microbiology, University
 College, Cork, College Road, Cork, Ireland

FEATURES
 source
 1..43075
 /organism="Streptococcus thermophilus temperate
 bacteriophage O1205"
 /strain="CNR21205"
 /db_xref="taxon:85154"
 complement(23..1102)
 /note="putative integrase"
 /codon_start=1
 /transl_table=11
 /evidence="not_experimental"
 /product="ORF1"
 /protein_id="AAC79517.1"
 /db_xref="GI:2444081"
 /translation="MASYRRKRNWERYINYYDSTGKRKPKTKGPRTKSEAIKAAAE
 MELKLDQNVNVEDITLYDFKWCWGVYKKTITKAYINISQRIELFFGDKKLK
 SITATEYQRLVNSYAKTAQDTRFNVHVKACIEMAHVEGYIKRNFCKFAKINAKNK
 GRDIETKFLVEYERLYETSKHPEYASYALYIISKTGIRFAECLGTVDDIKRDT
 GMLSVNKTWDYKNTGCMPTKTKSSIREIPLDDEFINEIDOLPPTDDGRLLSLSNNA
 VNKTLRKIVGREVRVHSLRHTYASYLIAHDIDLISVSQVLGHENLNTILEVIAHQLOE
 QKSRNDEKIKQMWTKCGQNALPHG"
 complement(1216..1686)
 CDS
 1216..1686
 /note="orf61.1"
 /codon_start=1
 /transl_table=11
 /protein_id="AAC79517.1"
 /db_xref="GI:2444081"
 /translation="MASYRRKRNWERYINYYDSTGKRKPKTKGPRTKSEAIKAAAE
 MELKLDQNVNVEDITLYDFKWCWGVYKKTITKAYINISQRIELFFGDKKLK
 SITATEYQRLVNSYAKTAQDTRFNVHVKACIEMAHVEGYIKRNFCKFAKINAKNK
 GRDIETKFLVEYERLYETSKHPEYASYALYIISKTGIRFAECLGTVDDIKRDT
 GMLSVNKTWDYKNTGCMPTKTKSSIREIPLDDEFINEIDOLPPTDDGRLLSLSNNA
 VNKTLRKIVGREVRVHSLRHTYASYLIAHDIDLISVSQVLGHENLNTILEVIAHQLOE
 QKSRNDEKIKQMWTKCGQNALPHG"
 complement(1216..1686)

CDS

/codon_start=1
/transl_table=11
/product="ORF2"
/protein_id="AAC79518.1"
/db_xref="GI:2444082"
/translation="MKDDVNRNPTYLKFTTFCGTEGGLILGFIPFPFWLLMFVGVA
LLFLPWEIKHPMTKEQBOLKIERAKAEERKSRNEFKAMAEARAVKCPYCKST
DVPEMVQORSFSGKAACTIMTGGVGALAGFAGKKGKKEWHCKNGCAVFTTK"
complement(1706..2092)
/codon_start=1
/transl_table=11
/product="ORF3"
/protein_id="AAC79519.1"
/db_xref="GI:2444083"
/translation="MKIDELLKEYKISLFIAPPAELWERSGFYPPDLRIYVNGSLSKQ
EREKVILHEIGHVNHDPKHKKYKRLQLQYENADRPMIRELLVYLKSTDIYDFNWRFA
AQVNISTWGEAMIQDEFKRIQQSVI"
complement(2082..2486)
/note="putative repressor"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="ORF4"
/protein_id="AAC79520.1"
/db_xref="GI:2444084"
/translation="MKGDVIMKENTAORLAQIMKERNLROVDILRQSEKFQKELGKL
GKSALSYVSGKSTPDODKLVLKSLTIGVSEAMLMGYDTNESTDVKKESIDLNLREK
VWMPDGPLSDDDVKKEIQIRLSIEVTGBED"
2635..2874
/codon_start=1
/transl_table=11
/product="ORF5"
/protein_id="AAC79521.1"
/db_xref="GI:2444085"
/translation="MIYKIDSKLNGKIVFVGTKKFAOSMNLSEKSISAKTNKKRC
WQSEISKACELLIQIPSEINSYFFRYKVQDIEHN"
2888..3169
/codon_start=1
/transl_table=11
/product="ORF6"
/protein_id="AAC79522.1"
/db_xref="GI:2444086"
/translation="MTDFPKPLADRFNLLTTVIADKSKAFDIDOTLPLILTAKQCOE
MLGIGNYTEFLRITNLDFGPKIDKGRGSQVRYPRDQVRQNNHWQDIA"
3230..3358
/codon_start=1
/transl_table=11
/product="ORF7"
/protein_id="AAC79523.1"
/db_xref="GI:2444087"
/translation="MKNLFKWILSKDEKQKPVWTPYEVNERKYEERHKREQLLMK"
3616..4089
/codon_start=1
/transl_table=11
/product="ORF8"
/protein_id="AAC79524.1"
/db_xref="GI:2444088"
/translation="MATLYELTDOLLEYNMVDDETGLDLEAIDWTTQYENKVEGY
VKYKISLEADIEARKNEKKRLDGLNKSQSKIDKLKTALAVSMAETGQTRVDTTLFKV
GFRKSEAVVNRNEKLPKEYQIATYKPKTKLLELLKSGKHIEGATLEERRNLNR"
4086..4787
/note="contains a putative ATP/GTP binding site between
amino acids 21-28"
/codon_start=1
/transl_table=11
/product="ORF9"
/protein_id="AAC79525.1"
/db_xref="GI:2444089"
/translation="MKIIRANDIORTKNWRIILYKAGLCKTSLIKNWPCKTIVLSLD
NSSKVGAGNVDIIDFDRHPTFETFLQADNLINKYENLVNDIISFSQSDWFIE
QGRSKNGINSELQHYSQWNTFLRVLTVISKPINITYAWEDTHEINLETGILTO
YVQIIRASVLNQLLGLTGVVGRIVVNAKTGARGLILEGSECTYAKNRLNDRACKIED
LKFQGLDGTKEPPE"

CDS

/codon_start=1
/transl_table=11
/product="ORF10"
/protein_id="AAC79526.1"
/db_xref="GI:2444090"
/translation="MELRNYSQSLVNDIKOSILRGNKRIMVQSPRSGKTVVMAHIAK
GATDKGNTILFFSHRKEINEOVNTFRKNGVDMNLVTIDSVTKIARNLDRIOEPSIIL
IDEAHVKAKTYLKIIEYNSIVLMFTGTPARLDGSGFDDIADDIVLGSVKRWLQDN
GNIAPEKYVAPSLDITNLLKRGGEFTKKSVDOTMKRVYICGDIIRHVEKLAGKQAIV
YTHSVEASVSNTFNBOGYTSAISGKTPPEVDRDMAQAFRDGKLTIMVNCLELFTG
IDLPNVDCIMLRPTQSILYLQFAMRALNPREGKTAIIIDHVCNVDRHGLPNADREN
SIGVNTKTKLKLGETPTTCDECYATFWSAERICPCMGCHENKPTKEEVEIIRIEL
EERRQEVASKEVETFTSDQCQTVLEELKFAQHKGICPGWVIYQOKENIMR"
6100..6535
/note="function unknown"
/codon_start=1
/transl_table=11
/product="ORF11"
/protein_id="AAC79527.1"
/db_xref="GI:2444091"
/translation="MFTIDYSQAKFEGSIKDGTYEVTIDAKQDATOGGADYLDIRFR
IRKDFQGEFQNNIIFIRYFAKKGDKGYPVASIMNIAKAAGIPDGTGFSSLEDYLNQLE
GKALKVTVKNEKSEWQGVENLNVKRLEVTDIPLEPVNTEISIDLPF"
6564..7373
/codon_start=1
/transl_table=11
/product="ORF12"
/protein_id="AAC79528.1"
/db_xref="GI:2444092"
/translation="MVDYAINYORWGSVIPISKNGKTPILISPADKPMPTENDIRRVV
RDNPANIALKTDTFFVIDVDMGDVGLNLNRNWEHARLIPPTLQAITPSGGRHIYL
KKDPNHPISQINIGIEGVDIKAHVNNTLVPNNSKGYIEMDTHVPGKGSITPEIYL
ALTKVLQMKPEPLSYESSFASDCVSRSTKTKLFLSILGFGDKGRNNALAEFVGG
LLLRGVDPEITYHLAKMANNTQEPGLKGEPTFKSMLDKEIRRIGTDND"
7360..8874
/note="putative primase"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="ORF13"
/protein_id="AAC79529.1"
/db_xref="GI:2444093"
/translation="MTTIDEDYREQFASLSPKKPSREGIKNKLKAYENDHFEKF
KEPNDRKRALPELAVAKGLNKTIVITLUNGKVAIYDPERGYQKDYRYAQLIY
VLEPTFNKCRNVFLFLSNMSREYENNMVMDPEYRDVRFILKNGIYDKRKKK
LLSFDYKFINFTIETELVENAPKPTINGWDVDSMLDMSGDSSELVELLMQVIAASL
NGNSYKRSIWLGVGNDGKGTFOQLISNLVGLKNVAPLKLNQFSERFGLAIIIEGKTV
IIGDVOAGIYVDESSNFNSVVTGEPVSIEKKGENPVLAQPKTVIOSTNAMPVFNK
SNGTYRIVLIIPKKTGINDDNALIKDDYINRKEVLEYLVKAINLDFDKFSEPKAT
QERMQEKEENNTVYKFLNEYLSDVSTRIPRELNDVYRSWCHGHNITIPKKSNEFK
ELAQNLPGWIKORQKPLDFNPTKDPVYVHDFENWMDENEAKKAAVVMVTO"
9128..9559
/codon_start=1
/transl_table=11
/product="ORF14"
/protein_id="AAC79530.1"
/db_xref="GI:2444094"
/translation="MEIQYLEINQEHFENENISNYIKDFSEAAATVIDVQCNAIPVHFE
KVGSDYWTDEDYGIKVVAFIKYEDNKATPEKKNLKEFFDKDKETKLFIDDPFDN
WATKGDVYVDTFRNCLFQQASMAVTDKQIVSIYKEWLQNH"
9601..9861
/codon_start=1
/transl_table=11
/product="ORF15"
/protein_id="AAC79531.1"
/db_xref="GI:2444095"
/translation="MVFIRNVGKVRMADGRWFDTGAPKGFCDLFGFRPDGOIFFIEVK
NEKRLRDLQKKFMKAMKRGALGVARSVKDAMDIVNEKRC"
9845..10081
/codon_start=1

CDS

4762..6093
/note="putative helicase"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="ORF10"
/protein_id="AAC79526.1"
/db_xref="GI:2444090"
/translation="MELRNYSQSLVNDIKOSILRGNKRIMVQSPRSGKTVVMAHIAK
GATDKGNTILFFSHRKEINEOVNTFRKNGVDMNLVTIDSVTKIARNLDRIOEPSIIL
IDEAHVKAKTYLKIIEYNSIVLMFTGTPARLDGSGFDDIADDIVLGSVKRWLQDN
GNIAPEKYVAPSLDITNLLKRGGEFTKKSVDOTMKRVYICGDIIRHVEKLAGKQAIV
YTHSVEASVSNTFNBOGYTSAISGKTPPEVDRDMAQAFRDGKLTIMVNCLELFTG
IDLPNVDCIMLRPTQSILYLQFAMRALNPREGKTAIIIDHVCNVDRHGLPNADREN
SIGVNTKTKLKLGETPTTCDECYATFWSAERICPCMGCHENKPTKEEVEIIRIEL
EERRQEVASKEVETFTSDQCQTVLEELKFAQHKGICPGWVIYQOKENIMR"
6100..6535
/note="function unknown"
/codon_start=1
/transl_table=11
/product="ORF11"
/protein_id="AAC79527.1"
/db_xref="GI:2444091"
/translation="MFTIDYSQAKFEGSIKDGTYEVTIDAKQDATOGGADYLDIRFR
IRKDFQGEFQNNIIFIRYFAKKGDKGYPVASIMNIAKAAGIPDGTGFSSLEDYLNQLE
GKALKVTVKNEKSEWQGVENLNVKRLEVTDIPLEPVNTEISIDLPF"
6564..7373
/codon_start=1
/transl_table=11
/product="ORF12"
/protein_id="AAC79528.1"
/db_xref="GI:2444092"
/translation="MVDYAINYORWGSVIPISKNGKTPILISPADKPMPTENDIRRVV
RDNPANIALKTDTFFVIDVDMGDVGLNLNRNWEHARLIPPTLQAITPSGGRHIYL
KKDPNHPISQINIGIEGVDIKAHVNNTLVPNNSKGYIEMDTHVPGKGSITPEIYL
ALTKVLQMKPEPLSYESSFASDCVSRSTKTKLFLSILGFGDKGRNNALAEFVGG
LLLRGVDPEITYHLAKMANNTQEPGLKGEPTFKSMLDKEIRRIGTDND"
7360..8874
/note="putative primase"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="ORF13"
/protein_id="AAC79529.1"
/db_xref="GI:2444093"
/translation="MTTIDEDYREQFASLSPKKPSREGIKNKLKAYENDHFEKF
KEPNDRKRALPELAVAKGLNKTIVITLUNGKVAIYDPERGYQKDYRYAQLIY
VLEPTFNKCRNVFLFLSNMSREYENNMVMDPEYRDVRFILKNGIYDKRKKK
LLSFDYKFINFTIETELVENAPKPTINGWDVDSMLDMSGDSSELVELLMQVIAASL
NGNSYKRSIWLGVGNDGKGTFOQLISNLVGLKNVAPLKLNQFSERFGLAIIIEGKTV
IIGDVOAGIYVDESSNFNSVVTGEPVSIEKKGENPVLAQPKTVIOSTNAMPVFNK
SNGTYRIVLIIPKKTGINDDNALIKDDYINRKEVLEYLVKAINLDFDKFSEPKAT
QERMQEKEENNTVYKFLNEYLSDVSTRIPRELNDVYRSWCHGHNITIPKKSNEFK
ELAQNLPGWIKORQKPLDFNPTKDPVYVHDFENWMDENEAKKAAVVMVTO"
9128..9559
/codon_start=1
/transl_table=11
/product="ORF14"
/protein_id="AAC79530.1"
/db_xref="GI:2444094"
/translation="MEIQYLEINQEHFENENISNYIKDFSEAAATVIDVQCNAIPVHFE
KVGSDYWTDEDYGIKVVAFIKYEDNKATPEKKNLKEFFDKDKETKLFIDDPFDN
WATKGDVYVDTFRNCLFQQASMAVTDKQIVSIYKEWLQNH"
9601..9861
/codon_start=1
/transl_table=11
/product="ORF15"
/protein_id="AAC79531.1"
/db_xref="GI:2444095"
/translation="MVFIRNVGKVRMADGRWFDTGAPKGFCDLFGFRPDGOIFFIEVK
NEKRLRDLQKKFMKAMKRGALGVARSVKDAMDIVNEKRC"
9845..10081
/codon_start=1

```

/transl_table=11
/product="ORF16"
/protein_id="AAC79532.1"
/db_xref="GI:2444096"
/transl_table="MRKNVKLVGIYAPGHYEHISVLEETQEFPSRWFWDHEDNEFI
SSKLGISTNKNRLITLQLPDELLARNGELCK"
10072..10245
/codon_start=1
/transl_table=11
/product="ORF17"

```

CDS

```

Query Match      73.3%; Score 17.6; DB 7; Length 43075;
Best Local Similarity 83.3%; Pred. NO. 1.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```
QY 1 tcgtcgaaagcttcgttaacgttcg 24
||||| ||||| ||||| ||||| |||||

```

```
Db 28112 TCGTCGAACCTTCGTATCGATAG 28089

```

RESULT 6

```

AC083943      AC083943      169696 bp      DNA      HTG      25-JUL-2001
LOCUS      Oryza sativa clone OSJNBa0044A10, WORKING DRAFT SEQUENCE, 4
DEFINITION      unordered pieces.
ACCESSION      AC083943
VERSION      AC083943.6 GI:15011678
KEYWORDS      HTG; HTGS-PHASE1; HTGS_DRAFT.
SOURCE      Oryza sativa.
ORGANISM      Oryza sativa

```

```

REFERENCE
AUTHORS      McCombie,W.R.
TITLE      Rice genomic sequence
JOURNAL      Unpublished
REFERENCE
AUTHORS      McCombie,W.R.
TITLE      Direct Submission
JOURNAL      Submitted (07-OCT-2000) Lita Annenberg Hazen Genome Center, Cold
Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor,
NY 11724, USA
COMMENT      On Jul 25, 2001 this sequence version replaced gi:14209714.
----- Genome Center
Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
Laboratory
Center code: CSHL
Web site: http://www.cshl.org/genseq
Contact: mcombie@cshl.org
----- Project Information
Center project name: OSJNBa0044A10
Center clone name: OSJNBa0044A10
-----

```

```

* NOTE: this is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 131206: contig of 131206 bp in length
* 131207 131392: gap of unknown length
* 131393 159122: contig of 27730 bp in length
* 159123 159308: gap of unknown length
* 159309 166084: contig of 6776 bp in length
* 166085 166269: gap of unknown length
* 166270 169696: contig of 3427 bp in length.

```

FEATURES

```

source
1..169696
/organism="Oryza sativa"
/db_xref="taxon:4530"

```

```
BASE COUNT      47992 a 36479 c 35957 g 48703 t      565 others
ORIGIN

```

```

Query Match      73.3%; Score 17.6; DB 2; Length 169696;
Best Local Similarity 83.3%; Pred. NO. 1.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```
QY 1 tcgtcgaaagcttcgttaacgttcg 24
||||| ||||| ||||| ||||| |||||

```

```
Db 17018 TCGTTAAGCGTCGTTACGTCG 17041

```

RESULT 7

```

AF041468      AF041468      121524 bp      DNA      circular      PLN      03-MAR-1999
LOCUS      Guillardia theta complete plastid genome.
DEFINITION      AF041468 X14171 X62349 X51511 X14504 X52158 X52912 X56806 M76547
ACCESSION      X62348 X21976 U81044 AF063017
KEYWORDS      AF041468.1 GI:3602932
VERSION      Guillardia theta.
SOURCE      Chloroplast Guillardia theta
ORGANISM      Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
REFERENCE
AUTHORS      1 (bases 47701 to 48415)
TITLE      Douglas,S.E. and Durnford,D.G.

```

```

The small subunit of ribulose-1,5-bisphosphate carboxylase is
plastid-encoded in the chlorophyll c-containing alga Cryptomonas
phi
Plant Mol. Biol. 13 (1), 13-20 (1989)
93357429
2 (bases 45872 to 47981)
Douglas,S.E., Durnford,D.G. and Morden,C.W.
Nucleotide sequence of the gene for the large subunit of
ribulose-1,5-bisphosphate carboxylase/oxygenase from the
chlorophyll c-containing Alga Cryptomonas F: evidence supporting
the polyphyletic origin of plastids
J. Phycol. 26, 500-508 (1990)
3 (bases 43739 to 44938)
Douglas,S.E. and Durnford,D.G.
Nucleotide sequence of the genes for ribosomal protein S4 and
trnA(Arg) from the chlorophyll c-containing alga Cryptomonas phi
Nucleic Acids Res. 18 (7), 1903 (1990)
90245597
4 (bases 18535 to 19351)
Douglas,S.E. and Durnford,D.G.
Sequence analysis of the plastid rDNA spacer region of the
chlorophyll c-containing alga Cryptomonas phi
DNA Seq. 1 (1), 55-62 (1990)
92119320
5 (bases 34539 to 35380)
Reith,M. and Douglas,S.
Localization of beta-phycoerythrin to the thylakoid lumen of
Cryptomonas phi does not involve a signal peptide
Plant Mol. Biol. 15 (4), 585-592 (1990)
91338697
6 (bases 110917 to 113854)
Douglas,S.E.
Unusual organization of a ribosomal protein operon in the plastid
genome of Cryptomonas phi: evolutionary considerations
Curr. Genet. 19 (4), 289-294 (1991)
91330343
7 (bases 40675 to 42376)
Douglas,S.E. and Turner,S.
Molecular evidence for the origin of plastids from a
cyanobacterium-like ancestor
J. Mol. Evol. 33 (3), 267-273 (1991)
92099311
8 (bases 96129 to 98906)
Wang,S.L. and Liu,X.Q.
The plastid genome of Cryptomonas phi encodes an hsp70-like
protein, a histone-like protein, and an acyl carrier protein

```

JOURNAL MEDLINE REFERENCE AUTHORS TITLE	Proc. Natl. Acad. Sci. U.S.A. 88 (23), 10783-10787 (1991) 92073372 9 (bases 106789 to 108216) Douglas,S.E. A scf homologue is found in the plastid genome of <i>Cryptomonas phi</i>	trna	/product="Clp protease ATP binding subunit" /protein_id="AAC35595.1" /db_xref="GI:3602934" /translation="MFERTKAIVKIMLAQBEARLGHNFVGTGEIILGLIGEGTGI AAKVLKSMVNLKRDARVEVEKTIIGRSGFVAVIEPTTPRAKRVLEUSLEAROLGHNY IGTEHLGLIREGEGVAARVLEALDITKVTQVIRLIGDPAEVSATNGTKGTP TLEFGSNTQKAAEGKDPVIGRQKEIBRVQILGRRTKNPILIGEPGVGKTAAE GLAQIRNRDVPDILEDKRVVTDIGLLVAGTKYRGEFERLKKIIDEIRVNNVILV IDVHTLICAGAAEADAANTILKPALARGEMOCIGATTLEEVKRIEKDSALERFQ PVMGEPSEETIEILYGLDRYKHHKLVISDEALSAAKFPADQIADRLDPKDAID LIDAGSRVRLMNSQLPAAREDKELREILKQKDEAVRQSDQFETAGQURDMEIKA QIAATAHKKKDEENTKEVSVTEEDIAQVAWTGIPVNMKTRSEKLLQMEETH GRIGODEAVAVSKAIRRVRVKNPPIASFIFSGPTGVGTETKALASYPFGS EAMVRLDMSEYMERHTVSKLSPGPGYVYGGGOLTESVRRRPVTVVLFDFIEKGH POVFNLLLOILEDGRTDSKGRVDFKNTLLILTSNVGSKVIEKGGVGLDFLSEDT ESQYGRKALVNEELKQYRPFELRLDEIIVFRQLTQDEVEGIAIEMLKVEYTRISE KGIOLEVTARFKTHLINEGNIYIYGARPLRRVNRULEDTLSEFPAEKIKEGDTAVV DVDDGKGVKVLGKLELVAN" 3316..3388 /gene="trnF(gaa)" /note="codon recognized: CUU" /product="trnA-Phe" /anticodon=(pos:3349..3351,aa:Phe) 3316..3388 /gene="trnF(gaa)" complement(3402..5201) /gene="dnaB" complement(3402..5201) /gene="dnaB" /codon_start=1 /transl_except=(pos:complement(5199..5201),aa:Met) /product="replication helicase subunit" /protein_id="AAC35596.1" /db_xref="GI:3602935"
gene		misc_feature	/translation="MKVTFYNLKAELHILQILVEKDFLQIYTKIDTQVYFQSHQ ILYRSINILHTRKVLNLPNLHLSSELVSETVINYLITLNDRONFRILNLSLQ LLDNEFIRRELQNSCAIVSLTFNFNLSVETLQKSNLISINTYKVKLSKISQL LLETILEIDKTNRSTHVLGTGDFDLHILVGLQKSDLIITAGRPMSGKTAFLMSLVNR VADIOSFPVIFESLMSKOLIVRLISNETNATSLRGENISISEWELNRAMTILS NLNYYDDNNLDVLDIQSKLASLOQYIGDGLIVDYLQLOQYKKSQRHIELSYI TRYLKIAKDNLPLVLSQLSNRVEFRNLKRPILSDLKESGCLSTVATQYFLKSDN VNFSLTSLKVSNNHYILSATLELLTVDLVATLQIYLYNNTEVDLFSNNVIFSSVINL VIVSTVEHVLTNGWKRIELTVDLVATLQIYLYNNTEVDLFSNNVIFSSVINL ICMNRINVTDFWIPKTNFVNALLVHNSDEQADVLFVLYRDSYVQHNLKQSNIDMC EVIYAKHRHGTIGMVNLIFNPNTVSPMNLKES" 3594..4073 /note="signal element intein; encodes putative intein spliced out of the replication helicase subunit" 5361..5433 /gene="trnW(cca)" /note="codon recognized: GGU" /product="trnA-Trp" /anticodon=(pos:5394..5396,aa:Trp) 5361..5433 /gene="trnW(cca)" 5588..6013 /gene="rpl11" 5588..6013 /gene="rpl11" /codon_start=1 /product="ribosomal protein L11" /protein_id="AAC35597.1" /db_xref="GI:3602936" /translation="MAKVVAVIKKALEAGKATPAPVGPALGQGVNIVMFCCKDYNA KQVEEIAQTKLPDLNTRRLESAIRIIEGTAKNMGIGVTD" 6033..6737 /gene="rpl11" 6033..6737 /gene="rpl11"
gene		gene	
CDS		CDS	
Query Match Best Local Similarity	72.5%; 94.7%;	Score 17.4; Pred. No. 2.4e+02;	DB 8; Length 121524;

```
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tcgtgaacgttcgtaac 19
|||||
Db 11476 TCGTCGACGTTCTTTAAC 11494

RESULT 8
AF007218 5905 bp DNA VRT 14-JUL-1997
LOCUS Tetraodon fluviatilis transcription factor snf5 (SNF5PF) gene,
DEFINITION complete cds.
ACCESSION AF007218
VERSION AF007218.1 GI:2253408
KEYWORDS
SOURCE Tetraodon fluviatilis.
ORGANISM Tetraodon fluviatilis.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 5905)
Yao,C.W., Leu,J.H. and Huang,C.J.
Genomic structure and sequence of SNF5 gene of puffer fish
(Tetraodon fluviatilis)
Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 5905)
AUTHORS Yao,C.W., Leu,J.H. and Huang,C.J.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-1997) Biological Chemistry, Academia Sinica, P.O.
Box 23-106, Taipei, Taiwan, Republic of China
FEATURES
source
1..5905
Location/Qualifiers
1..5905
/organism="Tetraodon fluviatilis"
/db_xref="taxon:47145"
join(2154..2325,3126..3236,3357..3488,3680..3817,
4101..4226,4299..4466,4617..4805,4877..5010,5110..5149)
/gene="SNF5PF"
/product="snf5"
2154..5149
/gene="SNF5PF"
join(2242..2325,3126..3236,3357..3488,3680..3817,
4101..4226,4299..4466,4617..4805,4877..5010,5110..5149)
/gene="SNF5PF"
/note="transcription factor"
/codon_start=1
/product="snf5"
/protein_id="AAB62935.1"
/db_xref="GI:2253408"
/translators="MALSFAFGQKPVKFOLEDGDFYMGSEVGNVLMFRGSLYKRY
PSLWRKLSVEERKKIVESHSDHGYTQATSVTLTKASEVEEILDGNDKRYKAVSIST
EPAYLREQAKRNSQWPTLPNSSHLDVAPCSTTINRSLRGDKKRTPLCFDDHD
PAVHNASQSEVLVPIRLDMEIEGQKRLDAFTWNNMKELMTPEMFAEILCDLIDLP
LAFVPAIPSAIROQIESPTDAILEQTDQDVITKLNHYGNTSLVDQFMDMSEKEN
SPKFAKLCSELGCGGEFVTIATYSTRGOLSHQKAYAFSENPLTVEIAIRNTGNA
DOMCPLELTIDAMEKKIRDQDNRIRRLANTAPGW"

BASE COUNT 1561 a 1389 c 1311 g 1644 t
ORIGIN

Query Match 71.7% Score 17.2; DB 5; Length 5905;
Best Local Similarity 86.4%; Pred. No. 3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 gtcgaacgttcgtaacgttcg 24
|||||
Db 4023 GTCGCTCGTTGGTTAACGTTGC 4044

RESULT 9
CNS07EFI 183317 bp DNA HTC
LOCUS Oryza sativa chromosome 12 clone OSJNba0009F13, *** SEQUENCING IN
DEFINITION

Query Match 71.7% Score 17.2; DB 2; Length 183317;
Best Local Similarity 86.4%; Pred. No. 3.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 tcgtgaacgttcgtaacgtt 22
|||||
Db 93471 TCGTTAAGCGTTCGTTAACGTT 93492

RESULT 10
AY024027 224 bp DNA PLN 07-FEB-2001
LOCUS Oryza sativa microsatellite MRG6352 containing (CGTT)X6, closest to
marker G132, genomic sequence.
DEFINITION
ACCESSION AY024027
VERSION AY024027.1 GI:12707243
KEYWORDS
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
Oryza.
1 (bases 1 to 224)
Tao,N., Barbazuk,W.B., Liu,J., Wu,K. and Barry,G.F.
Simple sequence repeats from Monsanto rice genomic sequences
Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 224)
AUTHORS Tao,N., Barbazuk,W.B., Liu,J., Wu,K. and Barry,G.F.

PROGRESS ***, in ordered pieces.
AL513003
AL513003.1 GI:12329144
HTG; HTGS_PHASE2.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 183317)
Salse,J., Cholsne,N., Orjeda,G., Regad,F., Lorieux,M., Cooke,R.,
Delseny,M., Robert,C., Brottier,P., Wincker,P., Cruaud,C.,
Artiguenave,F., Saurin,W., Salanoubat,M., Quetier,F. and
Weissenbach,J.
Oryza sativa chromosome 12 sequencing
Unpublished
2 (bases 1 to 183317)
Genoscope.
Direct Submission
Submitted (05-FEB-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence
Work on the sequence is in progress and the release of this data is
based on the understanding that the sequence may change as work
continue. The sequence may be contaminated with foreign sequence
from E.coli, yeast, vector, phage, etc.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1..183317
Location/Qualifiers
1..183317
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="12"
/clone="OSJNba0009F13"
/clone_lib="CUGI Nipponbare BAC"
52033 a 38000 c 38408 g 54873 t 3 others
ORIGIN

Query Match 71.7% Score 17.2; DB 2; Length 183317;
Best Local Similarity 86.4%; Pred. No. 3.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 tcgtgaacgttcgtaacgtt 22
|||||
Db 93471 TCGTTAAGCGTTCGTTAACGTT 93492

RESULT 10
AY024027 224 bp DNA PLN 07-FEB-2001
LOCUS Oryza sativa microsatellite MRG6352 containing (CGTT)X6, closest to
marker G132, genomic sequence.
DEFINITION
ACCESSION AY024027
VERSION AY024027.1 GI:12707243
KEYWORDS
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
Oryza.
1 (bases 1 to 224)
Tao,N., Barbazuk,W.B., Liu,J., Wu,K. and Barry,G.F.
Simple sequence repeats from Monsanto rice genomic sequences
Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 224)
AUTHORS Tao,N., Barbazuk,W.B., Liu,J., Wu,K. and Barry,G.F.
```

TITLE
JOURNAL

COMMENT
Direct Submission
Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh Blvd., Creve Coeur, MO 63167, USA
Derived from rice genomic sequences generated from the Monsanto Rice Genome Sequencing project. Please see <http://www.rice-research.org> for more information. The sequence data were produced primarily in the laboratories of Dr. Leroy Hood at the University of Washington in Seattle.

FEATURES
source 1. .224
/organism="Oryza sativa"
/db_xref="taxon:4530"
1. .224
/note="microsatellite MRG6352"
/rpt_type=tandem
/rpt_unit=cggt
BASE COUNT 30 a 57 c 42 g 95 t
ORIGIN

Query Match 70.8%; Score 17; DB 8; Length 224;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 acgttcgtaacgttcg 24
|||||
Db 50 ACGTCGTTACGTCG 66

RESULT 11
AP003370/c
LOCUS AP003370 101803 bp DNA HTG 08-MAR-2001
DEFINITION Oryza sativa chromosome 1 clone B1129H01, *** SEQUENCING IN PROGRESS ***, In ordered pieces.
ACCESSION AP003370
VERSION AP003370.1 GI:13365587
KEYWORDS HTG; HTGS-PHASE2
SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:B1129H01.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (sites)
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC clone:B1129H01
JOURNAL Published Only in DataBase (2001) In press
REFERENCE 2 (bases 1 to 101803)
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-2001) Takuji Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@abrr.affrc.go.jp, URL:<http://rgp.dna.affrc.go.jp/>, Tel:81-298-38-7441, Fax:81-298-38-7468)
NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source 1. .101803
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="B1129H01"
BASE COUNT 29451 a 21109 c 21813 g 29380 t 50 others

ORIGIN

Query Match 70.8%; Score 17; DB 2; Length 101803;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 acgttcgtaacgttcg 24
|||||
Db 96635 ACGTCGTTACGTCG 96619

RESULT 12
AP003987
LOCUS AP003987 104718 bp DNA HTG 02-AUG-2001
DEFINITION Oryza sativa chromosome 2 clone OJ1057_A12, *** SEQUENCING IN PROGRESS ***, In ordered pieces.
ACCESSION AP003987
VERSION AP003987.1 GI:15076790
KEYWORDS HTG; HTGS-PHASE2
SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:OJ1057_A12.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 104718)
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC clone:OJ1057_A12
JOURNAL Published Only in DataBase (2001) In press
REFERENCE 2 (bases 1 to 104718)
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-2001) Takuji Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@abrr.affrc.go.jp, URL:<http://rgp.dna.affrc.go.jp/>, Tel:81-298-38-7441, Fax:81-298-38-7468)
NOTE: The nucleotide sequence of this BAC clone was generated by combining Monsanto and RGP-Japan sequencing data.
NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source 1. .104718
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="2"
/clone="OJ1057_A12"
BASE COUNT 31132 a 21354 c 21209 g 30922 t 101 others
ORIGIN

Query Match 70.8%; Score 17; DB 2; Length 104718;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 acgttcgtaacgttcg 24
|||||
Db 79837 ACGTCGTTACGTCG 79853

RESULT 13
AP003196
LOCUS AP003196 149498 bp DNA HTG 21-FEB-2001

DEFINITION Oryza sativa chromosome 1 clone B1008C01, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION AP003196
VERSION AP003196.1 GI:13027226
KEYWORDS HTG: HTGS_PHASE2.
SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:B1008C01.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (sites)
REFERENCE
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
clone:B1008C01
JOURNAL Published Only in DataBase (2001) In press
REFERENCE
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@agr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
Location/Qualifiers
1..149498
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="B1008C01"
BASE COUNT 41722 a 33641 c 31241 g 42644 t 250 others
ORIGIN
Query Match 70.8%; Score 17; DB 2; Length 149498;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 8 acgttcgttaacgttcg 24
|||||
Db 25666 ACGTTCGTTACGTTTCG 25682
RESULT 14
AP003074/c
LOCUS AP003074 150379 bp DNA PLN 29-MAR-2001
DEFINITION Oryza sativa genomic DNA, chromosome 1, BAC clone:OSJNBa0004G10.
ACCESSION AP003074
VERSION AP003074.2 GI:13486822
KEYWORDS
SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:OSJNBa0004G10.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (sites)
REFERENCE
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
clone:OSJNBa0004G10
JOURNAL Published Only in DataBase (2001) In press
REFERENCE
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Direct Submission
Submitted (10-JAN-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@agr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Mar 28, 2001 this sequence version replaced gi:12082349.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI NonRedundant protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DDBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from -21M13 to M13rev of the BAC
clone. This sequence of OSJNBa0004G10 clone has an overlap with
P0034C11 (DDBJ: AP002865) clone at the position 78,934 to 150,379
of 3' end. The sequence of this clone ends at the position 71,446
of P0034C11. Detailed information on overlap and assembly quality
together with annotation of this entry is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.
Location/Qualifiers
1..150379
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="OSJNBa0004G10"
join(3892..3894,4077..4235,4302..4379,4831..4989)
/gene="OSJNBa0004G10.1"
join(3892..3894,4077..4235,4302..4379,4831..4989)
/gene="OSJNBa0004G10.1"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB40054.1"
/db_xref="GI:13486823"
/translation="MTYRPMCNFIEWDMENPQNNGTFRAYPRNETRSDYLRRKDEHER
RIAALAEWHVIRSVRHRTRGCVCPNTVDDDFVCGNPRMGFTRWIDNVNTPSYHQC
KITESETQVEYQRLKHENAMHSDRPRRG"
complement(join(6294..6572,7027..7047))
/gene="OSJNBa0004G10.2"
complement(join(6294..6572,7027..7047))
/gene="OSJNBa0004G10.2"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB40055.1"
/db_xref="GI:13486824"
/translation="MIPSRVQPSLVDEDDGGGGRGQRAQOVGDGRGGGSGVDGGRTPR
WRWRRGTTATTAVEDGGGGGGGGRVRORRGRMRALRWCAIPLGDISPSILV"
join(9066..9225,10554..10996)
/gene="OSJNBa0004G10.3"
join(9066..9225,10554..10996)
/gene="OSJNBa0004G10.3"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB40056.1"
/db_xref="GI:13486825"
/translation="MEENEVEDDNIPIFYQVGFEGNOTGEERDADGNDVADDLGQI
LDQVKEDCETLSPISLSLSVAAGRLVADADVVAAEHDSAPRRRGTTTTRDATT
PPRHGTAAAPRLRTTTPPPRAAAAPRAALPPRCTPHRCRAHAAAPRRHSAAT
PALCRCHVNERINVNERMNVNERNVREELNVNERA"
complement(join(13279..13346,13549..14688,14843..14926,
15165..15308,15344..15407))
/gene="OSJNBa0004G10.4"
complement(join(13279..13346,13549..14688,14843..14926,


```
15165..15308,15344..15407))
/genes="OSJNBa0004G10.4"
/notes="hypothetical protein
similar to Oryza sativa OSJNBa0049B20.1"
/codon_start=1
/db_xref="GI:13486826"
/translation="MPDWLGACVDAGOTPVRCARCPCAPSSRVVPPFAHGNRTSFRRGA
LASIDPSPARLSMRLADSRTRYSPRNYPARAIIHGNSIPSRXYKPRHNGGTPEPSEA
PMKRVADTRAKRKQKQKGRKWLILACVPHVYRGDYFEGAGMDNVIKAALAPRATRLV
VHRSAVTRRKTIDDIHGFVAGADCHGRLLYASQGPPEPVIDLGYRGPLGHHGFPK
AFYICDLKHKSTPLDHGFPILHFGNAGLVAITKTIFHVAIDLHPTVGSQKAAALIYF
SYPEIATHWVYPPDRPWGGVGVVHKTIIVMWVDSLGLSCDISARRQDLRFVPL
CPDCLPCTPDLDRKRCVFRGDDLRVMGITHREYDVKPVVSMWTLVDVQDAGTWRLD
COPLVKDIWDEGYKATKLPREIPTVAFIHPGLGNVAFEMRSRLGVLNLYTRKYLE
WQFFAMLPMPRYHSSRFVRACPNGLIMHALCEISTRNYSFPMHG"
complement(19438..20550)
/genes="OSJNBa0004G10.5"
complement(19438..20550)
/genes="OSJNBa0004G10.5"
/notes="hypothetical protein
similar to Oryza sativa OSJNBa0049B20.1"
/codon_start=1
/protein_id="BAB40058.1"
/db_xref="GI:13486827"
/translation="MOMREAAKGMWVILVSPHVAHDEFHPPGTGLRNLFGAPPEAS
RTVPRIAPDKADINYPYLAADERHGRLLLYATQDPDPEPRPALDAFYLRPLGVH
HGFAKAYFICDTTITREASRLPDHPDHPFALYPGNVGLLCYSISFYVAELQAPASGTA
TULLYKSDSDAWDELSYPPHDPHMGWGMVSHQRLMWVDSLGLLTCDDVYGGDP
PLHLYVPPDSELPACTPDLKRCRCVGSAGRLRYQIDDEPDGDPPIVMYMLTLDDE
AGEMFGDCAASFAVWDDAYATKPLQVPAVALIHPTGPGDVVYFELRSRIFAVDV
RARRLLEWFFEMLHPMPRYHSSQFVRANMLPTLFQ"
complement(22493..22509,23300..23459,23521..23766,
24024..24080))
/genes="OSJNBa0004G10.6"
complement(Join(22493..22509,23300..23459,23521..23766,
24024..24080))
/genes="OSJNBa0004G10.6"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB40059.1"
/db_xref="GI:13486828"
/translation="MSCRAGTFCGAPPPVWPMLPLPLRAAGHRRGRHQAAPRCIKAS
AGNMRYVIGHGPDPDEQVYTIWTLHDRVHARWRREYEVFVETWDCGRKRKSEMERKG
QLMLGAGGSGVRLRDPLTKTPSVGTLPPFVRVWEVPPSLRTPVPPVQHDNLN"
Join(26535..27327,27406..27422)
/genes="OSJNBa0004G10.7"
Join(26535..27327,27406..27422)
/genes="OSJNBa0004G10.7"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB40060.1"
/db_xref="GI:13486829"
/translation="MANPVPCPSAAAYGPGCGDVGCGSSSHNVNWAHRVHRHNGVYRL
CYFVCLHHRDAFCVCLDFPDGAPDDFYNPVSCCGGPEVPAARAAHACLADLPDS
YFVCPACAAAGERTFYAPSSGAPHALAEVLLVAARLHAHESVRAAAAAREQAERL
IRAAARARRMDVACVLEAEARDAEQAAAPPVLTKTTPKNSAARRSDDK
PLKINSIQPALAFAAAAAAASSTPLSTPSPAGERKPKQGRVLI"
complement(Join(28267..28383,28533..28659,28780..28926,
29466..30857,31383..31453))
/genes="OSJNBa0004G10.8"
complement(Join(28267..28383,28533..28659,28780..28926,
29466..30857,31383..31453))
/genes="OSJNBa0004G10.8"
/notes="hypothetical protein
similar to Oryza sativa OSJNBa0049B20.1"
/codon_start=1
/protein_id="BAB40061.1"
/db_xref="GI:13486830"
/translation="MIRSRGSTGTVPACAAVNRDRCKPTGSSSPFPVHPPTPAA
TDHNSAAAAAASEVRSSRRMEHSHPLFDGAAPKFRDERMAAPREATWILACVP
VSSSDSEFAGDHLAFDWRDPPGVSLTLRQSDSVFVSPAPRDFCPDRDDHPYVAV
DSAGGLLLRGARRSAHDFGFGVALGDPAPSCNTDGGYILCHATLRMAYLYPPCSDEY
```

```
RLLCAGNVGMIIRRTADRDHPRIRLAELQIESNGIHRATLLRYSHELGGWASTKVNYP
PGRSSCGSVLVHACMLAWVDLSFGLLTCDVFAAKPDMRFVPLPPECCKLPYSDDADH
AKHRCVNVSDGELAFVQIHVDYDTAAGRGAPSTIMISWTLQOSDAGEESVSLRHRVR
VDEIWDHVTYRKTMPPRRVPLALLHPKELGVVFFVITSRNSMFAVDLVTLVLEEC
KKYKMPQLPTMYHSSHRVAMELPHSICRGEDEDTGTIACVNLONLTDLIDLNFSSD
KADELLSTTGRLEINFRFQELRNATAPFKYLSFVIVKATDAYEALCLMRNFVSHVSMO
GNIIVPRVSAESIISGLTLCYLTQPVFVNMWADDIVRDGTLLSSNRIFDG"
complement(Join(32717..32742,33198..33313,33434..33558,
33802..33841,34125..35525,36069..36148,36260..36304))
/genes="OSJNBa0004G10.9"
complement(Join(32717..32742,33198..33313,33434..33558,
33802..33841,34125..35525,36069..36148,36260..36304))
/genes="OSJNBa0004G10.9"
/notes="hypothetical protein
similar to Oryza sativa OSJNBa0049B20.1"
/codon_start=1
/protein_id="BAB40062.1"
/db_xref="GI:13486831"
/translation="MVSSSSAATRCGPRRITAKIDTVSRFGSSSINSFAIYICADGK
LIDGSPFPVHTPPPIAATDHSNAAAAAASEVRSSRRMEHSHPLTLRQSDSVFVSP
RMAAPREATWILACVPVSSSDGDFEAGDHLAFDWRDPPGVSLTLRQSDSVFVSP
APRDFCPDRDDHPVYVAVDSAGLLLRGARRSHNDRGPGVALGDPAPSCANDGGYIL
CHATLRMAYLYPPCSDIYRLLYAGNVGMIRRTAADGHPRIALLAEQIESNGIHRAT
LLRYSELGLGGWASTKVNYPGRSSCGDGVIVHAGMLWVDLSFGLLTCDFAAARP
DMRFVPLPPECCKLPYSDDADHAKHRCVNVSDGELAFVQIHVDYTAAGRGAPSTIMIS
WTLQOSDAGEESVSLRHRVDEIWDHVTYRKTMPPRRVPLALLHPKELGVVFFFO
ITSRNSMFAVDLVTLVLECKYKMPQLPTMYHSSHRVAMELPHSICRGEDEDTG
TSYNVNLTFDHSVNLTDLNLNFSRDADELLSTTGRLEINFRFQELRNATAPFKYL
SFVIVKATDAYEASCLMRDFVSHVSMGDSNIVPRVSEPEYDDDD"
Join(36644..36865,37413..37473,37564..37949)
/genes="OSJNBa0004G10.10"
Join(36644..36865,37413..37473,37564..37949)
/genes="OSJNBa0004G10.10"
/notes="hypothetical protein
similar to Oryza sativa OSJNBa0049B20.2"
/codon_start=1
```

Query Match 70.8%; Score 17; DB 8; Length 150379;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 acgttgtaacgttcg 24
|||||

Db 11078 ACGTTCGTTAACGTTTCG 11062

RESULT 15
AP003211/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AP003211 150641 bp DNA HTG 21-FEB-2001
Oryza sativa chromosome 1 clone OSJNBa0011P19, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
AP003211 GI:13027241
HTG; HTGS_PHASE2.
Oryza sativa (cultivar:Nipponbare) DNA, clone:OSJNBa0011P19.
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (sites)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
clone:OSJNBa0011P19
Published Only in DataBase (2001) In press
2 (bases 1 to 150641)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasak@agr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
NOTE: It currently consists of 1 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES
source

Location/Qualifiers
1. .150641
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="OSJNBa0011P19"

BASE COUNT 41696 a 33905 c 33240 g 41500 t 300 others
ORIGIN

Query Match 70.8%; Score 17; DB 2; Length 150641;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 acgttcgttaacgttcg 24
|||||

Db 148877 ACGTTCGTAAACGTTCG 148861

Search completed: March 11, 2002, 20:56:28
Job time: 4346 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 20:59:44 ; Search time 237.21 Seconds
(without alignments)
90.559 Million cell updates/sec

Title: US-09-802-518-9
Perfect score: 24
Sequence: 1 tcgtcgaaacgttcgttaacgttcg 24

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 42862619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
C 1	16.4	68.3	2958	13	AAQ25266	Corynebacterium gl
C 2	16.4	68.3	2958	13	AAQ25263	C.glutamium-deriv
C 3	16.4	68.3	2958	13	AAQ25265	C.glutamium-deriv
C 4	16	66.7	4142	20	AAQ84319	Stearth virus nucl
C 5	16	66.7	4414	21	AAQ47124	Arabidopsis thalia
C 6	16	66.7	13732	21	AAQ81484	N. meningitidis pa
C 7	16	66.7	34980	21	AAQ21611	Neisseria meningit
C 8	15.8	65.8	506	20	AAV86097	EST clone D310. H
C 9	15.8	65.8	9619	20	AAQ13134	Enterococcus faeca
C 10	15.6	65.0	555	22	AAQ03845	Complement of huma
C 11	15.6	65.0	608	18	AAV75360	Staphylococcus aur

C 12	15.6	65.0	1436	21	AAQ64784	Lemon alcohol acyl
C 13	15.6	65.0	3600	21	AAA98970	S. aureus mprF DNA
C 14	15.4	64.2	20	21	AAA96274	Sequence of a stab
C 15	15.4	64.2	20	21	AAA96280	Aspergillus niger
C 16	15	62.5	468	21	AAF11680	Arabidopsis thalia
C 17	15	62.5	506	21	AAQ44327	C glutamicum codin
C 18	15	62.5	738	22	AAH65227	S. epidermidis ope
C 19	15	62.5	1089	22	AAH53863	S. epidermidis ope
C 20	15	62.5	1188	21	AAQ50733	Yeast type II topo
C 21	15	62.5	1320	20	AAQ78298	R. capsulatus hema
C 22	15	62.5	1479	21	AAQ39143	Arabidopsis thalia
C 23	15	62.5	1574	21	AAQ45383	Arabidopsis thalia
C 24	15	62.5	1579	21	AAQ33969	Arabidopsis thalia
C 25	15	62.5	1584	21	AAQ49134	Arabidopsis thalia
C 26	15	62.5	1635	20	AAQ95510	Nucleic acid seque
C 27	15	62.5	1767	21	AAQ32958	Arabidopsis thalia
C 28	15	62.5	1769	21	AAQ49141	Arabidopsis thalia
C 29	15	62.5	1947	22	AAH81338	Escherichia coli p
C 30	15	62.5	3603	22	AAH54113	S. epidermidis gen
C 31	15	62.5	3757	22	AAH54828	S. epidermidis gen
C 32	15	62.5	3816	22	AAH54850	S. epidermidis gen
C 33	15	62.5	7424	22	AAQ28521	Genomic fragment #
C 34	15	62.5	133719	21	AAQ64754	Macaca mulatta rha
C 35	15	62.5	349980	22	AAH64966	C glutamicum codin
C 36	14.8	61.7	936	21	AAQ42756	Arabidopsis thalia
C 37	14.8	61.7	1154	21	AAQ45380	Arabidopsis thalia
C 38	14.8	61.7	1180	20	AAQ13531	Enterococcus faeca
C 39	14.8	61.7	1197	21	AAQ39350	Arabidopsis thalia
C 40	14.8	61.7	2958	13	AAQ25264	C.glutamium-deriv
C 41	14.8	61.7	4529	19	AAV16865	Genomic DNA sequen
C 42	14.6	60.8	318	21	AAQ75078	Human ORFX ORF633
C 43	14.6	60.8	524	21	AAQ35552	Arabidopsis thalia
C 44	14.6	60.8	643	21	AAQ08385	Fusarium venenatum
C 45	14.6	60.8	828	22	AAH67608	C glutamicum codin

ALIGNMENTS

RESULT 1
AAQ25266/c
ID AAQ25266 standard; DNA; 2958 BP.
XX
AC AAQ25266;
XX
DT 27-NOV-1992 (first entry)
XX
DE Corynebacterium glutamicum ORI sequence.

XX
Origin of replication: plasmid pHM1519; DS gene; ATCC 13058;
KW 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase; mutant;
KW aromatic amino acid synthesis; retinol inhibition; aspartate kinase; ss.

XX
Corynebacterium glutamicum.
XX
PN FR2667875-A.
XX
PD 17-APR-1992.
XX
PF 15-OCT-1991; 91FR-0012701.
XX
PR 15-OCT-1990; 90JP-0273348.
XX
PR 19-JUN-1991; 91JP-0245291.
XX
(AJIN) AJINOMOTO KK.
XX
PI Kojima H, Matsui H, Nakamatsu T, Satoh K, Sugimoto M;
XX Tanaka A;
XX
DR WPI; 1992-185928/23.
XX
PT New Corynebacterium plasmids with heat sensitive origin of
replication - inhibited and expelled at high temp. used to

PT transform microorganisms by homologous recombination, esp. for
 PT increasing amino acid yields in fermentation
 XX
 PS Example 1; Page 14; 29pp; French.
 XX

CC The origin of replication from plasmid pHM1519, naturally present in
 CC C.glutamicum, was inserted into shuttle vector pH4. The resulting
 CC plasmid, pHK4, can replicate in E.coli and Corynebacterium. pH4 was
 CC treated with hydroxylamine and the products were used to transform
 CC Brevibacterium lactofermentum FERM P-7559. Three kanamycin-resistant
 CC strains were unable to develop at high temperatures. These strains
 CC contained the plasmids pHSC4, pHSC22 and pHSC23, respectively.
 CC See AAQ25263-5.
 XX
 SQ Sequence 2958 BP; 700 A; 854 C; 835 G; 569 T; 0 other;

Query Match 68.3%; Score 16.4; DB 13; Length 2958;
 Best Local Similarity 94.4%; Pred. No. 49;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 cgaacgttcgttaacgtt 22
 |||||
 DB 432 CGAACGTTTCGTTAAAGTT 415

RESULT 2
 AAQ25263/c
 ID AAQ25263 standard; DNA; 2958 BP.
 XX
 AC AAQ25263;
 XX
 DT 27-NOV-1992 (first entry)
 DE C.glutamicum-derived temp-sensitive ORI present in pHSC4.
 XX
 KW Origin of replication; plasmid pHSC4; DS gene; ATCC 13058;
 KW 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase; mutant;
 KW aromatic amino acid synthesis; retroinhibition; aspartate kinase; ss.
 XX
 OS Corynebacterium glutamicum.
 XX
 FH Key Location/Qualifiers
 FT mutation 1543
 FT /*tag= a
 FT /*note= "Wild-type G substituted by A"
 FT mutation 1546
 FT /*tag= b
 FT /*note= "Wild-type G substituted by A"
 XX
 XX FR2667875-A.
 XX
 XX 17-APR-1992.
 XX
 XX 15-OCT-1991; 91FR-0012701.
 XX
 XX 15-OCT-1990; 90JP-0273348.
 PR 19-JUN-1991; 91JP-0245291.
 XX
 PA (AJIN) AJINOMOTO KK.
 XX
 XX Kojima H, Matsui H, Nakamatsu T, Satoh K, Sugimoto M;
 PI Tanaka A;
 XX
 XX WPI; 1992-185928/23.
 XX
 XX New Corynebacterium plasmids with heat sensitive origin of
 PT replication - inhibited and expelled at high temp. used to
 PT transform microorganisms by homologous recombination, esp. for
 PT increasing amino acid yields in fermentation
 XX
 PS Claim 3; Page 21; 29pp; French.
 XX

CC The origin of replication from plasmid pHM1519, naturally present in
 CC C.glutamicum, was inserted into shuttle vector pH4. The resulting
 CC plasmid, pHK4, can replicate in E.coli and Corynebacterium. pH4 was
 CC treated with hydroxylamine and the products were used to transform
 CC Brevibacterium lactofermentum FERM P-7559. Three kanamycin-resistant
 CC strains were unable to develop at high temperatures. These strains
 CC contained the plasmids pHSC4, pHSC22 and pHSC23, respectively. The
 CC mutant ORI sequences were determined. They are each capable of
 CC autonomous replication and maintenance in Corynebacterium but
 CC replication is inhibited and the plasmids are expelled from their
 CC hosts when cultured at 31-37 deg.C. The plasmids contg. the
 CC temperature-sensitive sequences are used to produce transformant
 CC Corynebacterium spp. by homologous recombination. A mutant gene
 CC (such as 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase or
 CC aspartate kinase mutated to eliminate retroinhibition) is ligated to
 CC the plasmid contg. the mutant ORI. The wild-type gene undergoes
 CC homologous recombination and is expelled from the cell as part of
 CC the plasmid at a temperature of 31-37 deg.C.
 CC See also AAQ25264-6.
 XX
 SQ Sequence 2958 BP; 702 A; 854 C; 833 G; 569 T; 0 other;

Query Match 68.3%; Score 16.4; DB 13; Length 2958;
 Best Local Similarity 94.4%; Pred. No. 49;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 cgaacgttcgttaacgtt 22
 |||||
 DB 432 CGAACGTTTCGTTAAAGTT 415

RESULT 3
 AAQ25265/c
 ID AAQ25265 standard; DNA; 2958 BP.
 XX
 AC AAQ25265;
 XX
 DT 27-NOV-1992 (first entry)
 DE C.glutamicum-derived temp-sensitive ORI present in pHSC23.
 XX
 KW Origin of replication; plasmid pHSC23; DS gene; ATCC 13058;
 KW 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase; mutant;
 KW aromatic amino acid synthesis; retroinhibition; aspartate kinase; ss.
 XX
 OS Corynebacterium glutamicum.
 XX
 FH Key Location/Qualifiers
 FT mutation 756
 FT /*tag= a
 FT /*note= "Wild-type G replaced by A"
 FT mutation 1561
 FT /*tag= b
 FT /*note= "Wild-type G replaced by A"
 FT mutation 1668
 FT /*tag= c
 FT /*note= "Wild-type C replaced by T"
 FT mutation 1685
 FT /*tag= d
 FT /*note= "Wild-type C replaced by T"
 XX
 XX FR2667875-A.
 XX
 XX 17-APR-1992.
 XX
 XX 15-OCT-1991; 91FR-0012701.
 XX
 XX 15-OCT-1990; 90JP-0273348.
 PR 19-JUN-1991; 91JP-0245291.
 XX
 PA (AJIN) AJINOMOTO KK.
 XX

PI Kojima H, Matsui H, Nakamatsu T, Satoh K, Sugimoto M;
 PI Tanaka A;
 DR WPI; 1992-185928/23.
 XX
 XX New Corynebacterium plasmids with heat sensitive origin of
 PT replication - inhibited and expelled at high temp. used to
 PT transform microorganisms by homologous recombination, esp. for
 PT increasing amino acid yields in fermentation
 XX
 XX Claim 5; Page 24; 29pp; French.
 XX
 CC The origin of replication from plasmid pHM1519, naturally present in
 CC C-glutamicum, was inserted into shuttle vector pH4. The resulting
 CC plasmid, pHK4, can replicate in E.coli and Corynebacterium. pH4 was
 CC treated with hydroxylamine and the products were used to transform
 CC Brevibacterium lactofermentum FERM p-7559. Three kanamycin-resistant
 CC strains were unable to develop at high temperatures. These strains
 CC contained the plasmids pHSC4, pHSC22 and pHSC23, respectively. The
 CC mutant ORI sequences were determined. They are each capable of
 CC autonomous replication and maintenance in Corynebacterium but
 CC replication is inhibited and the plasmids are expelled from their
 CC hosts when cultured at 31-37 deg.C. The plasmids contg. the
 CC temperature-sensitive sequences are used to produce transformant
 CC Corynebacterium spp. by homologous recombination. A mutant gene
 CC (such as 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase or
 CC aspartate kinase mutated to eliminate retroinhibition) is ligated to
 CC the plasmid contg. the mutant ORI. The wild-type gene undergoes
 CC homologous recombination and is expelled from the cell as part of
 CC the plasmid at a temperature of 31-37 deg.C.
 CC See AAQ25263-6.
 XX
 XX Sequence 2958 BP; 702 A; 851 C; 833 G; 572 T; 0 other;
 SQ

Query Match 68.3%; Score 16.4; DB 13; Length 2958;
 Best Local Similarity 94.4%; Pred. No. 49;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 cgaacgttcgttaacgtt 22
 |||||
 Db 432 CGAAGTTCGTAAAGTT 415

RESULT 4
 AAX84319/C
 ID AAX84319 standard; DNA; 4142 BP.
 AC
 XX AAX84319;
 XX
 XX 08-SEP-1999 (first entry)
 DT
 XX
 DE Stealth virus nucleic acid clone, SEQ ID NO: 11;
 XX
 XX Stealth virus; detection; diagnosis; infection; ss.
 KW
 XX Stealth virus.
 OS
 XX WO9934019-A1.
 FN
 XX 08-JUL-1999.
 PD
 XX 30-DEC-1998; 98WO-US27744.
 PF
 XX 30-DEC-1997; 97US-0001184.
 PR
 XX (MARTIN) MARTIN W J.
 PA
 XX Martin WJ;
 PI
 XX WPI; 1999-405521/34.
 DR
 XX Novel strains of stealth virus
 PT

XX Claim 19; Page 42-44; 95pp; English.
 XX
 CC This sequence represents a Stealth virus nucleic acid clone. The
 CC invention relates to a method of detecting and characterising a stealth
 CC virus by reacting a sample suspected of containing a stealth virus with a
 CC probe from a known stealth virus and sequencing the resultant isolated
 CC nucleotide. The method comprises the steps of: (a) isolating DNA or RNA
 CC from a sample suspected of containing a stealth virus, e.g. a culture of
 CC cells showing a viral cytopathic effect; (b) testing the reactivity of
 CC the isolated DNA or RNA with a molecular probe that contains at least 18
 CC or more contiguous nucleotides identical to sequence previously
 CC identified from a stealth virus; and, optionally (c) sequencing the
 CC isolated DNA or RNA molecules that react with the probe. The method is
 CC used to detect stealth virus in a biological product, food or in the
 CC environment. The method is also used to evaluate agents for their
 CC inhibitory or stimulatory effects on stealth virus replication and to
 CC determine capacity of the virus to recombine with and potentially alter
 CC the nucleic acid sequences of a cell or bacterium.
 XX
 SQ Sequence 4142 BP; 1029 A; 964 C; 1082 G; 1065 T; 2 other;
 Query Match 66.7%; Score 16; DB 20; Length 4142;
 Best Local Similarity 79.2%; Pred. No. 77;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 tcgtcgaaacgttcgttaacgttcg 24
 |||||
 Db 289 TCGTCGAACGCTCTTCACTATCG 266

RESULT 5
 AAC47124
 ID AAC47124 standard; DNA; 4414 BP.
 AC
 XX AAC47124;
 XX
 XX 18-OCT-2000 (first entry)
 DT
 XX
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 52650.
 DE
 XX Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 XX EP1033405-A2.
 PN
 XX 06-SEP-2000.
 PD
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 PF
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126284.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.

PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 66.7%; Score 16; DB 21; Length 4414;
 Best Local Similarity 79.2%; Pred. No. 76;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 tcgtcgaaacgttcgttaacgttcg 24
 ||||| | ||||| |||||
 Db 230 tcgtccagcttcgttgacgtcg 253

RESULT 6
 AAA81484/C
 ID AAA81484 standard; DNA; 13732 BP.

XX AC AAA81484;
 XX DT 04-DEC-2000 (first entry)
 XX DE N. meningitidis partial DNA sequence gnm_32 SEQ ID NO:32.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KW Meningococcus B; MenB; ds.

XX OS Neisseria meningitidis.
 XX PN WO200022430-A2.
 XX PD 20-APR-2000.
 XX PF 08-OCT-1999; 99WO-US23573.

XX PR 09-OCT-1998; 99US-0103794.
 XX PR 30-APR-1999; 99US-0132068.
 XX PA (CHIR) CHIRON CORP.
 XX PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
 PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Pizza M;
 XX WPI; 2000-318079/27.

XX PT Isolated nucleotide sequences of Neisseria meningitidis which can be
 PT used in the diagnosis and treatment of N. meningitidis infection and
 PT other Neisserial infections, for example, N.gonorrhoea -
 XX Claim 7; Page 589-593; 1760pp; English.
 XX CC The present invention describes methods of obtaining immunogenic
 CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
 CC represent specifically claimed Neisseria meningitidis genomic DNA
 CC sequences: AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
 CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
 CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
 CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
 CC sequences, which are all used in the exemplification of the present

CC invention. The nucleic acid sequences, protein sequences, and antibodies
 CC against them, can be used in the manufacture of a composition. The
 CC composition can be used as a medicament (or in the manufacture of a
 CC medicament) for treating, preventing or diagnosing infection due to
 CC Neisserial bacteria. For example, some of the identified proteins could
 CC be components of vaccines against Meningococcus B; against all serotypes;
 CC and/or against all pathogenic Neisseriae. Identification of sequences
 CC from the bacterium will also facilitate production of biological probes,
 CC particularly organism-specific probes. Attempts to make efficacious
 CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
 CC Multivalent vaccines have also been tried but none have successfully
 CC overcome antigenic variability. The provision of further, complete
 CC sequences may provide an opportunity to identify secreted or surface
 CC exposed proteins that may be presumed targets for the immune system and
 CC which are not antigenically variable or at least more conserved than
 CC other more variable regions.

XX SQ Sequence 13732 BP; 3488 A; 3771 C; 3357 G; 3116 T; 0 other;

Query Match 66.7%; Score 16; DB 21; Length 13732;
 Best Local Similarity 79.2%; Pred. No. 75;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 tcgtcgaaacgttcgttaacgttcg 24
 ||||| | ||||| |||||
 Db 13293 TCGCGAACATCCGTTCCGGTTCG 13270

RESULT 7
 AAF21611/C
 ID AAF21611 standard; DNA; 349980 BP.

XX AC AAF21611;
 XX DT 13-MAR-2001 (first entry)
 XX DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:112.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
 KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
 XX ds.
 XX OS Neisseria meningitidis.
 XX PN WO2000066791-A1.
 XX PD 09-NOV-2000.
 XX PF 08-MAR-2000; 2000WO-US05928.
 XX PR 30-APR-1999; 99US-0132068.
 XX PR 08-OCT-1999; 99WO-US23573.
 XX PR 28-FEB-2000; 2000GB-0004695.
 XX PA (CHIR) CHIRON CORP.
 XX PA (GENO-) INST GENOMIC RES.

XX PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Maignani V;
 PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
 PI Frazer CM, Grandi G;
 XX WPI; 2000-647603/62.
 XX PT Neisseria meningitidis B full length genome sequence and open reading
 PT frames are used to detect, treat and prevent Neisserial infections -
 XX Claim 7; Appendix A; 692pp; English.
 XX CC The present invention describes the full length genome of
 CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
 CC to AAF21613 represent fragments of the NMB genomic sequence, as the
 CC sequence was too long to go in a record on its own it was split into 8

CC can be used in vaccines to prevent or attenuate an Enterococcal
CC infection.
XX
SQ Sequence 9619 BP; 2909 A; 1604 C; 2033 G; 3062 T; 11 other;

Query Match 65.8%; Score 15.8; DB 20; Length 9619;
Best Local Similarity 89.5%; Pred. No. 94;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 cgtcgaaacgttcgtaacy 20
||| ||||| ||||| |||||
Db 9133 CGTTGAACGTCCTTTAACG 9115

RESULT 10
AAD03845/C
ID AAD03845 standard; DNA; 555 BP.

XX AAD03845;

XX 31-JUL-2001 (first entry)

XX Complement of human B cell maturation protein (BCMA) DNA.

XX Human; A Proliferation Inducing Ligand Receptor; APRIL-R; cytostatic;
KW gene therapy; cancer; nephrotropic; renal disorder; autoimmune disease;
KW carcinoma; lung; colon; breast; prostate; Grave's disease; hypertension;
KW systemic lupus erythematosus; SLE; inflammation; cardiovascular disease;
KW B-cell lympho-proliferative disorder; BCM; immunosuppressive disease;
KW organ transplantation; HIV; human immunodeficiency virus; TNF;
KW tumour necrosis factor; BCMA; B cell maturation protein; ds.

XX Homo sapiens.

XX WO200124811-A1.

XX 12-APR-2001.

XX 05-OCT-2000; 2000WO-US27579.

XX 06-OCT-1999; 99US-0157933.

XX 11-FEB-2000; 2000US-0181807.

XX 30-JUN-2000; 2000US-0215688.

XX (BIOJ) BIOGEN INC.

XX (APOT-) APOTEC R & D SA.

XX Schneider P, Thompson J, Cachero T, Ambrose C, Rennert P;

XX WPI; 2001-266242/27.

XX Treating a mammal for a condition associated with undesired cell

XX proliferation such as cancer or carcinoma, comprises administering a

XX composition comprising A Proliferation Inducing Ligand Receptor

XX (APRIL-R) antagonist

XX Disclosure; Fig 3A; 85pp; English.

XX The invention relates to a method of treating a mammal for a condition

XX associated with undesired cell proliferation such as cancer or

XX carcinoma. The method involves administering a composition comprising

XX A Proliferation Inducing Ligand Receptor (APRIL-R) also referred as

XX B cell maturation protein (BCM or BCMA) antagonist that antagonises the

XX interaction between APRIL and its cognate receptor(s). This method is

XX useful for treating undesired cell proliferation such as cancer or

CC altering an immune response involving a signalling pathway between
CC APRIL-R and its ligand. APRIL-R DNA is also useful in gene therapy.
CC The present sequence is a complement of human APRIL-R also referred
XX as BCMA DNA.

XX Sequence 555 BP; 152 A; 132 C; 115 G; 156 T; 0 other;

Query Match 65.0%; Score 15.6; DB 22; Length 555;
Best Local Similarity 81.8%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 tcgtcgaaacgttcgtaacgtt 22
||||| ||| || |||||
Db 513 TCGTCGACCGTCCGAGACGTT 492

RESULT 11

AAV75360

ID AAV75360 standard; DNA; 608 BP.

XX AAV75360;

XX 16-MAR-1999 (first entry)

XX Staphylococcus aureus contig SEQ ID #1049.

XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.

XX Staphylococcus aureus.

XX EP786519-A2.

XX 30-JUL-1997.

XX 07-JAN-1997; 97EP-0100117.

XX 05-JAN-1996; 96US-0009861.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;

XX Rosen CA;

XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
XX stored on computer readable medium and used in the production of
XX anti-S.aureus vaccines

XX Claim 1; Page 1809; 327lpp; English.

XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
XX of the invention. The DNA sequences are recorded on a computer readable
XX medium, preferably selected from a floppy or hard disk, random access
XX memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
XX the S.aureus DNA sequences allows putative functions to be assigned so
XX that protein-encoding or regulatory regions of commercial, therapeutic or
XX industrial importance can be obtained. Specifically, sequences which are
XX likely to encode antigens have been identified and these polypeptides can
XX be used in a vaccine composition against S.aureus infection. The
XX polypeptides can also be used in a kit for the immunodetection of
XX S.aureus in a sample. S.aureus is implicated in numerous human diseases,
XX including cellulitis, eyelid infections, food poisoning, osteomyelitis,
XX skin and surgical wound infections, scalded skin syndrome, toxic shock
XX syndrome, etc. Organisms transformed with the DNA sequences can be used
XX for recombinant production of the polypeptides. The new DNA sequences
XX (and their fragments) are useful as primers or probes for isolating
XX homologues of any of the S.aureus DNA sequences contained on the
XX computer readable medium.

The present sequence represents a stabilised oligonucleotide which has antitumour activity. The oligonucleotide comprises an octamer motif of the type 5'-purine-purine-CG-pyrimidine-pyrimidine-X-X-3', where the pair X-X is AT, AA, CT or TT. The oligonucleotides are immunostimulatory, and are not toxic. They may be adapted for use in animals or humans. The stabilised oligonucleotides are used for treating tumours, of any type and any degree of anaplasia, particularly human tumours in the peripheral or central nervous systems, specifically

Qy 7 aacgttcgtaacgttc 23
||||| |||||
pb 3 aacgttccttaacgttc 19

Search completed: March 11, 2002, 20:59:57
Job time: 4180 sec

SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC 13058
US-08-408-188A-1

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 cgaacgttcgtaacgtt 22
|||||
Db 432 CGAACGTTCTGTTAAAGTT 415

RESULT 2
US-08-408-188A-2/c
Sequence 2, Application US/08408188A
Patent No. 5616480
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: KOJIMA, Hiroyuki
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
APPLICANT: SATOH, Katsuaki
APPLICANT: NAKAMATSU, Tsuyoshi
TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,188A
FILING DATE: 22-MAR-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/182,361
FILING DATE: 18-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/774,846
FILING DATE: 11-OCT-1991
APPLICATION NUMBER: JP 2-273348
FILING DATE: 15-OCT-1990
APPLICATION NUMBER: JP 3-245291
FILING DATE: 19-JUN-1991

ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC 13058
US-08-408-188A-2

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 cgaacgttcgtaacgtt 22
|||||
Db 432 CGAACGTTCTGTTAAAGTT 415

RESULT 3
US-08-408-188A-3/c
Sequence 3, Application US/08408188A
Patent No. 5616480
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: KOJIMA, Hiroyuki
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
APPLICANT: SATOH, Katsuaki
APPLICANT: NAKAMATSU, Tsuyoshi
TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,188A
FILING DATE: 22-MAR-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/182,361
FILING DATE: 18-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/774,846
FILING DATE: 11-OCT-1991
APPLICATION NUMBER: JP 2-273348
FILING DATE: 15-OCT-1990
APPLICATION NUMBER: JP 3-245291
FILING DATE: 19-JUN-1991

ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: Corynebacterium glutamicum
;; STRAIN: ATCC 13058
US-08-408-188A-3

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. NO. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 cgaacgttcgttaacgtt 22
|||||
Db 432 CGAACGTCGTTAAAGTT 415

RESULT 4
US-08-408-188A-4/c
; Sequence 4, Application US/08408188A
; Patent No. 5616480
; GENERAL INFORMATION:
; APPLICANT: SUGIMOTO, Masakazu
; APPLICANT: KOJIMA, Hiroyuki
; APPLICANT: TANAKA, Akiko
; APPLICANT: MATSUI, Hiroshi
; APPLICANT: SATOH, Katsuaki
; APPLICANT: NAKAMATSU, Tsuyoshi
; TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,188A
FILING DATE: 22-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/182,361
FILING DATE: 18-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/774,846
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-273348
FILING DATE: 15-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-245291
FILING DATE: 19-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

;; ORIGINAL SOURCE:
;; ORGANISM: Corynebacterium glutamicum
;; STRAIN: ATCC 13058
US-08-408-188A-4

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. NO. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 cgaacgttcgttaacgtt 22
|||||
Db 432 CGAACGTCGTTAAAGTT 415

RESULT 5
US-08-766-488-1/c
; Sequence 1, Application US/08766488
; Patent No. 5756347
; GENERAL INFORMATION:
; APPLICANT: SUGIMOTO, Masakazu
; APPLICANT: KOJIMA, Hiroyuki
; APPLICANT: TANAKA, Akiko
; APPLICANT: MATSUI, Hiroshi
; APPLICANT: SATOH, Katsuaki
; APPLICANT: NAKAMATSU, Tsuyoshi
; TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,488
FILING DATE: 13-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/408,188
FILING DATE: 22-MAR-1995
APPLICATION NUMBER: US 08/182,361
FILING DATE: 18-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/774,846
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-273348
FILING DATE: 15-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-245291
FILING DATE: 19-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:
; ORGANISM: Corynebacterium glutamicum
; STRAIN: ATCC 13058
US-08-766-488-1

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 cgaacgttcgttaacggtt 22
|||||
Db 432 CGAACGTCGTAAAGTT 415

RESULT 6

US-08-766-488-2/c
; Sequence 2, Application US/08766488
; Patent No. 5756347
; GENERAL INFORMATION:
; APPLICANT: SUGIMOTO, Masakazu
; APPLICANT: KOJIMA, Hiroyuki
; APPLICANT: TANAKA, Akiko
; APPLICANT: MATSUI, Hiroshi
; APPLICANT: SATOH, Katsuaki
; APPLICANT: NAKAMATSU, Tsuyoshi
; TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/766,488
; FILING DATE: 13-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/408,188
; FILING DATE: 22-MAR-1995
; APPLICATION NUMBER: US 08/182,361
; FILING DATE: 18-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/774,846
; FILING DATE: 11-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-273348
; FILING DATE: 15-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-245291
; FILING DATE: 19-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2959 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:
; ORGANISM: Corynebacterium glutamicum
; STRAIN: ATCC 13058
US-08-766-488-2

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 cgaacgttcgttaacggtt 22
|||||
Db 432 CGAACGTCGTAAAGTT 415

RESULT 7

US-08-766-488-3/c
; Sequence 3, Application US/08766488
; Patent No. 5756347
; GENERAL INFORMATION:
; APPLICANT: SUGIMOTO, Masakazu
; APPLICANT: KOJIMA, Hiroyuki
; APPLICANT: TANAKA, Akiko
; APPLICANT: MATSUI, Hiroshi
; APPLICANT: SATOH, Katsuaki
; APPLICANT: NAKAMATSU, Tsuyoshi
; TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/766,488
; FILING DATE: 13-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/408,188
; FILING DATE: 22-MAR-1995
; APPLICATION NUMBER: US 08/182,361
; FILING DATE: 18-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/774,846
; FILING DATE: 11-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-273348
; FILING DATE: 15-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-245291
; FILING DATE: 19-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2959 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

;; ORIGINAL SOURCE:
;; ORGANISM: Corynebacterium glutamicum
;; STRAIN: ATCC 13038
US-08-766-488-3

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 cgaacgttcgttaacgtt 22
|||||
Db 432 CGAACGTCGTTAAAGTT 415

RESULT 8
US-08-766-488-4/c
; Sequence 4, Application US/08766488
; Patent No. 5756347
; GENERAL INFORMATION:
; APPLICANT: SUGIMOTO, Masakazu
; APPLICANT: KOJIMA, Hiroyuki
; APPLICANT: TANAKA, Akiko
; APPLICANT: MATSUI, Hiroshi
; APPLICANT: SATOH, Katsuaki
; APPLICANT: NAKANATSU, Tsuyoshi
; TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/766,488
; FILING DATE: 13-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/408,188
; FILING DATE: 22-MAR-1995
; APPLICATION NUMBER: US 08/182,361
; FILING DATE: 18-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/774,846
; FILING DATE: 11-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-273348
; FILING DATE: 15-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-245291
; FILING DATE: 19-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2959 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

;; ORIGINAL SOURCE:
;; ORGANISM: Corynebacterium glutamicum
;; STRAIN: ATCC 13058
US-08-766-488-4

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 cgaacgttcgttaacgtt 22
|||||
Db 432 CGAACGTCGTTAAAGTT 415

RESULT 9
US-08-565-386-4
; Sequence 4, Application US/08565386
; Patent No. 5741697
; GENERAL INFORMATION:
; APPLICANT: Bavoll, Patrik M.
; APPLICANT: Hsia, Ru-ching
; TITLE OF INVENTION: BACTERIOPHAGE OF CHLAMYDIA PSITTACI
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Roches
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/565,386
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 176/60040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-565-386-4
Query Match 61.7%; Score 14.8; DB 1; Length 447;
Best Local Similarity 88.9%; Pred. No. 65;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 cgaacgttcgttaacgtt 22
|||||
Db 248 CGAACGTCGTTAAAGTT 265

RESULT 10
US-08-565-386-1
; Sequence 1, Application US/08565386
; Patent No. 5741697
; GENERAL INFORMATION:
; APPLICANT: Bavoll, Patrik M.
; APPLICANT: Hsia, Ru-ching
; TITLE OF INVENTION: BACTERIOPHAGE OF CHLAMYDIA PSITTACI
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Roches
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/565,386
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 176/60040
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4529 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-565-386-1

Query Match 61.7%; Score 14.8; DB 1; Length 4529;
Best Local Similarity 88.9%; Pred. No. 78;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 cgaacgttcgttaacgtt 22
||||| ||||| |||||
Db 2731 CGAAGTTCGTGAACGTT 2748

RESULT 11
US-08-844-154-1
Sequence 1, Application US/08844154
Patent No. 5827708
GENERAL INFORMATION:
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: NO. 5827708el Compounds
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/844,154
FILING DATE: 18-APR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9607993.4
FILING DATE: 18-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmli, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31457-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1303 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-844-154-1

Query Match 60.8%; Score 14.6; DB 1; Length 1303;
Best Local Similarity 81.0%; Pred. No. 89;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 tcgaacgttcgttaacgttcg 24
||||| ||||| |||||
Db 1256 TCGAACGTATGTTAACCTTCG 1276

RESULT 12
US-09-126-192A-1
Sequence 1, Application US/09126192A
Patent No. 6040163
GENERAL INFORMATION:
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 6040163el Compounds
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/09/126,192A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/844,154
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gimmli, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31457-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1303 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-126-192A-1

Query Match 60.8%; Score 14.6; DB 3; Length 1303;
Best Local Similarity 81.0%; Pred. No. 89;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 tcgaacgttcgttaacgttcg 24
||||| ||||| |||||
Db 1256 TCGAACGTATGTTAACCTTCG 1276

```
RESULT 13
US-08-902-585-1
; Sequence 1, Application US/08902585
; Patent No. 5834276
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 5834276el Asparaginyl tRNA Synthetase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,585
; FILING DATE: 29-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,154
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: 9607993.4
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ginni, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31457-3/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1344 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-902-585-1

Query Match 60.8%; Score 14.6; DB 2; Length 1344;
Best Local Similarity 81.0%; Pred. No. 89;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 tcgaacgttcgttaacgttcg 24
||||||| | |||||
Db 1256 TCGAACGTATGTTACCTTCG 1276

RESULT 14
US-08-961-083-221/c
; Sequence 221, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Chol et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
US-08-961-083-221/c
; Sequence 221, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Chol et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
```

```
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 221:
SEQUENCE CHARACTERISTICS:
LENGTH: 1777 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-221

Query Match 60.8%; Score 14.6; DB 3; Length 1777;
Best Local Similarity 81.0%; Pred. No. 91;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 cgtcgaaacttcgttaacgtt 22
||||||| | |||||
Db 65 CGTCGAATGTACGTGAACCTT 45

RESULT 15
US-08-961-083-219/c
; Sequence 219, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Chol et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 1879 base pairs
TYPE: nucleic acid
```

; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-083-219

Query Match 60.88; Score 14.6; DB 3; Length 1879;
Best Local Similarity 81.0%; Pred. No. 92;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 2 cgtcgaaacgttcgttaacgtt 22
||||||| || ||||| |||||
Db 1865 CGTCGAATGTACGTGACCTT 1845

Search completed: March 11, 2002, 21:01:48
Job time: 4235 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 20:30:00 ; Search time 2230.08 Seconds
(without alignments)
115.645 Million cell updates/sec

Title: US-09-802-518-9

Perfect score: 24

Sequence: 1 tgcgcgaacgttcgttaacgttcg 24

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- EST:*
- 1: em_estfun:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estom:*
 - 5: em_estpl:*
 - 6: em_estba:*
 - 7: em_estro:*
 - 8: em_estov:*
 - 9: em_htc:*
 - 10: gb_estl:*
 - 11: gb_est2:*
 - 12: gb_htc:*
 - 13: gb_gss:*
 - 14: em_gss_fun:*
 - 15: em_gss_hum:*
 - 16: em_gss_inv:*
 - 17: em_gss_pln:*
 - 18: em_gss_pro:*
 - 19: em_gss_rod:*
 - 20: em_gss_vit:*
 - 21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	17.6	73.3	531	10	AV403813
2	17.6	73.3	571	10	AL118705
3	17.6	73.3	609	13	AQ623639
4	17.2	71.7	794	13	CNS02XJ3
5	17.2	71.7	2228	11	BF860303
6	16.6	69.2	360	11	D37092
7	16.6	69.2	379	11	C12646
8	16.6	69.2	460	10	AI531538
c 9	16.6	69.2	464	13	AQ860867
10	16.6	69.2	475	10	AW942684
c 11	16.6	69.2	522	10	AA517098
12	16.6	69.2	532	11	BG638934

13	16.6	69.2	537	10	AI976120	AI976120	EST270714
c 14	16.6	69.2	552	11	BG795168	UTSW_SM2E	
c 15	16.6	69.2	576	11	BG799488	fm97f06.y	
c 16	16.6	69.2	586	11	BG799472	fm97d08.y	
c 17	16.6	69.2	598	11	BF815578	MR2-C1012	
18	16.6	69.2	672	10	AW017352	EST272774	
c 19	16.6	69.2	696	11	BG641369	SD12691.5	
c 20	16.6	69.2	713	13	CNS03K6W	Tetraodon	
c 21	16.6	69.2	799	11	BG638867	LD38308.3	
c 22	16.6	69.2	893	10	BE413349	MCS012.B0	
c 23	16.6	69.2	947	11	BG823392	602726796	
c 24	16.6	69.2	1095	11	BF686639	602143569	
c 25	16.6	69.2	1892	11	BG847877	1024019E0	
c 26	16.4	68.3	454	11	BG639770	SD05702.3	
c 27	16.4	68.3	1019	11	BF682153	602117263	
c 28	16.2	67.5	546	10	AI437242	fb39d01.y	
c 29	16.2	67.5	548	13	AQ968186	LERIW64TR	
c 30	16.2	67.5	611	11	BG891803	OO_L4_O3G	
c 31	16.2	67.5	624	13	AQ968185	LERIW64TF	
c 32	16.2	67.5	639	11	BG304106	FL30d09.x	
c 33	16.2	67.5	644	10	AW154870	EST290263	
c 34	16.2	67.5	644	11	BF677194	602087262	
c 35	16.2	67.5	701	13	AQ362660	nbxb0050K	
c 36	16.2	67.5	870	13	CNS01I8T	AL15422	Anopheles
c 37	16.2	67.5	913	13	CNS01VO4	AL169357	Tetraodon
c 38	16.2	67.5	987	11	BG176913	602313482	
c 39	16.2	67.5	1051	13	CNS04PA3	AL301044	Tetraodon
c 40	16.2	67.5	1083	13	CNS051V8	AL317357	Tetraodon
c 41	16	66.7	360	11	D67466	DG7466	CELK052H7F
c 42	16	66.7	368	10	AU036374	AU036374	
c 43	16	66.7	384	10	AA123623	MBAFCE8D1	
c 44	16	66.7	399	13	AQ852283	MAJFVL1	
c 45	16	66.7	457	13	TA364A06P	AL494140	T. brucei

ALIGNMENTS

RESULT 1

AV403813/c

LOCUS

DEFINITION

AV403813 Bombyx mori pheromone gland Shuko x Ryuaku newly-eclosed adult Bombyx mori cDNA clone pg--0086 T3, mRNA sequence.

ACCESSION

AV403813

VERSION

AV403813.1 GI:6907901

KEYWORDS

EST.

SOURCE

domestic silkworm.

ORGANISM

Bombyx mori

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.

REFERENCE

1 (bases 1 to 531)

Mita, K., Morimyo, M., Shimada, T., Okano, K. and Maeda, S.

PROJECT=Silkworm Genome Program in MAFF, and Research for the Future Program in JSPS. see 'SilkBase', <http://www.ab.a.u-tokyo.ac.jp/silkbase/>, for whole ESTdb. Method: uni-directional, sequence direction: sequenced from T3 primer (5' -> 3').

AUTHORS

Bombyx mori cDNA

UNPUBLISHED (2000)

CONTACT: Mita K

Genome Research Group

National Institute of Radiological Sciences

Anagawa 4-9-1, Inage, Chiba 263-8555, Japan

Email: kmita@nirs.go.jp

FEATURES

source

1. 531

/organism="Bombyx mori"

/strain="Shuko x Ryuaku"

/db_xref="taxon:7091"

/clone="pg--0086"

/clone.lib="Bombyx mori pheromone gland Shuko x Ryuaku newly-eclosed adult"

/sex="female"
/tissue_type="pheromone gland"
/dev_stage="newly-eclosed adult"
119 a 134 c 148 g 130 t

BASE COUNT
ORIGIN

Query Match 73.3%; Score 17.6; DB 10; Length 531;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 tcgtcgacgttcgttaacgttcg 24
|||||
Db 485 TCGTCGAACCGCGTTAAACGTCG 462

RESULT 2
AL118705

LOCUS AL118705 571 bp mRNA EST 29-FEB-2000
DEFINITION DKFZp761l0510_r1 761 (synonym: hamy2) Homo sapiens cDNA clone
ACCESSION DKFZp761l0510 5', mRNA sequence.

VERSION AL118705
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE EST (Blum, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Blum H

MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by LMU (Ludwig Maximilians University,
Munich/Germany) within the cDNA sequencing consortium of the German
Genome Project.

No sl sequence available.

This clone (DKFZp761l0510) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..571
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp761l0510"
/clone_lib="761 (synonym: hamy2)"
/tissue_type="amygdala"
/dev_stage="adult"
/lab_host="DH10B"
/note="vector: pSport1; site_1: NotI; Site_2: SalI"

BASE COUNT
ORIGIN

Query Match 73.3%; Score 17.6; DB 10; Length 571;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 tcgtcgacgttcgttaacgttcg 24
|||||
Db 452 TCGACGGACGTTCTGTAACCGTCG 475

RESULT 3

AQ623639

LOCUS AQ623639 609 bp DNA GSS 16-JUN-1999

DEFINITION HS_5377_A2_F05_SF6E RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=953 Col=10 Row=K, DNA sequence.

ACCESSION AQ623639.1 GI:5086119
VERSION A0623639.1
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 609)
MAHAIAS.G.G., WALLACE, J.C., SMITH, K., SWARTZELL, S., HOLZMAN, T.,
KELLER, A., SHAKER, R., FURLONG, J., YOUNG, J., ZHAO, S., ADAMS, M.D. and
HOOD, L.

Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589

TITLE

JOURNAL High Throughput Sequencing Center
MEDLINE University of Washington
COMMENT Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 953 Row: K Column: 10
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 609.

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)

or from Resear h Genetics (info@resgen.com). BAC end Web Server:

http://www.htsc.washington.edu

Plate: 953 Row: K Column: 10

Seq primer: SP6

Class: BAC ends

High quality sequence stop: 609.

Location/Qualifiers

1..609

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Plate=953 Col=10 Row=K"

/clone_lib="RPCI-11 Human Male BAC Library"

/sex="male"

/note="vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;

Male blood DNA was isolated from one randomly chosen donor

and partially digested with a combination of EcoRI and

EcoRI Methylase. Size selected DNA was cloned into the

pBACe3.6 vector at EcoRI sites"

BASE COUNT 185 a 124 c 162 g 125 t 13 others

ORIGIN

Query Match 73.3%; Score 17.6; DB 13; Length 609;

Best Local Similarity 83.3%; Pred. No. 1e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 tcgtcgacgttcgttaacgttcg 24
|||||
Db 335 TCGTCGAACGTTCTGTAACGTCG 358

RESULT 4

CNS02XJ3

LOCUS

DEFINITION CNS02XJ3 794 bp DNA GSS 15-MAY-2000
Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
178F10 of library G from Tetraodon nigroviridis, genomic survey
sequence.

AL218424

AL218424.1 GI:7877243

GSS; genome survey sequence.

Tetraodon nigroviridis.

Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;

Tetraodontidae; Tetraodon.

1 (bases 1 to 794)

REFERENCE

AUTHORS

Roest-Crolius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bonneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.

TITLE

Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis

JOURNAL

REFERENCE

AUTHORS

REFERENCE

TITLE

Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence

JOURNAL

REFERENCE

AUTHORS

REFERENCE

TITLE

REFERENCE

JOURNAL

REFERENCE

COMMENT

REFERENCE

FEATURES

source

1..794

/organism="Tetraodon nigroviridis"

/db_xref="taxon:99883"

/clone="178F10"

/clone_lib="C"

/notes="Genoscope sequence ID : COAG178DC05SP1-end :

PUC-ori"

BASE COUNT 162 a 232 c 239 g 161 t

ORIGIN

Query Match

Best Local Similarity 71.7%; Score 17.2; DB 13; Length 794;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy

3 gtcgaacgttcgttaacgttcg 24

||||| ||||| ||||| ||||| |||||

Db 11 GTCGCTGTTGGTTAAGTTCG 32

RESULT 5

BF860303

LOCUS

DEFINITION

963016B12.x1 C. reinhardtii CC-1690, Stress condition I, normalized

, Lambda zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

BF860303

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..2228

/organism="Chlamydomonas reinhardtii"

/strain="CC-1690 wild type mt+ 21gr"

/db_xref="taxon:3055"

AUTHORS

/clone_lib="C. reinhardtii CC-1690, Stress condition I,
normalized, Lambda Zap II"

TITLE

/notes="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2:
XhoI; This library, constructed by John Davies and Jeffrey

JOURNAL

REFERENCE

AUTHORS

REFERENCE

TITLE

Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence

JOURNAL

REFERENCE

AUTHORS

REFERENCE

TITLE

REFERENCE

JOURNAL

REFERENCE

COMMENT

REFERENCE

FEATURES

source

1..360

/organism="Caenorhabditis elegans"

/strain="CB1489 him-8(el489)"

/db_xref="taxon:6239"

/clone="yk40a3"

/clone_lib="Yuji Kohara unpublished cDNA"

/sex="hermaphrodite, male"

/tissue_type="whole animal"

/dev_stage="varied"

BASE COUNT 95 a 76 c 73 g 115 t 1 others

ORIGIN

Query Match

Best Local Similarity 69.2%; Score 16.6; DB 11; Length 360;

Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy

2 cgtcgaacgttcgttaacgttcg 24

||||| ||||| ||||| ||||| |||||

Db 204 CGTCGAAATTCCTTTACGATCG 226

McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,
1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr
, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). POLYA mRNA was
purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into lambda Zap II
(Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with ExAssist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 633 a 427 c 543 g 470 t 135 others

ORIGIN

Query Match

Best Local Similarity 71.7%; Score 17.2; DB 11; Length 2228;

Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy

1 tcgtcgaacgttcgttaacgttcg 24

||||| ||||| ||||| ||||| |||||

Db 1744 TCGTCGAATGTTTCGATAGNCTCG 1767

RESULT 6

D37092

LOCUS

DEFINITION

CELK040E3F Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA

clone_yk40e3 5', mRNA sequence.

D37092

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..360

/organism="Caenorhabditis elegans"

/strain="CB1489 him-8(el489)"

/db_xref="taxon:6239"

/clone="yk40a3"

/clone_lib="Yuji Kohara unpublished cDNA"

/sex="hermaphrodite, male"

/tissue_type="whole animal"

/dev_stage="varied"

BASE COUNT 95 a 76 c 73 g 115 t 1 others

ORIGIN

Query Match

Best Local Similarity 82.6%; Score 16.6; DB 11; Length 360;

Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy

2 cgtcgaacgttcgttaacgttcg 24

||||| ||||| ||||| ||||| |||||

Db 204 CGTCGAAATTCCTTTACGATCG 226

RESULT 7

LOCUS C12646 379 bp mRNA EST 28-DEC-1998
 DEFINITION C12646 Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA
 clone yk153h4 5', mRNA sequence.
 ACCESSION C12646
 VERSION C12646.1 GI:1560199
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans.
 ORGANISM Caenorhabditis elegans.
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Peloderinae; Caenorhabditis.
 REFERENCE 1 (bases 1 to 379)
 AUTHORS Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
 M., Miyata,A. and Nishigaki,A.
 TITLE Expression map of the C.elegans genome
 JOURNAL Unpublished (1996)
 COMMENT Contact: Yuji Kohara
 Genome Biology Lab.
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykohara@lab.nig.ac.jp.
 Location/Qualifiers
 1..379
 /organism="Caenorhabditis elegans"
 /strain="CEL489 him-8(el489)"
 /db_xref="taxon:6239"
 /clones="yk153h4"
 /clone_lib="Yuji Kohara unpublished cDNA"
 /sex="hermaphrodite, male"
 /tissue_type="whole animal"
 /dev_stage="varied"

FEATURES source

BASE COUNT 95 a 90 c 94 g 92 t 8 others
 ORIGIN
 Query Match 69.2%; Score 16.6; DB 11; Length 379;
 Best Local Similarity 79.2%; Pred. No.3e+02; 5; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 tcgtcgacgttcgttaacgttcg 24
 ||||| ||||| ||||| ||||| |||||
 Db 115 TCGTGGACATTCGNCACATTCG 138

RESULT 8

LOCUS AI531538 460 bp mRNA EST 19-APR-2001
 DEFINITION SD02623.3prime SD Drosophila melanogaster Schneider L2 cell culture
 POT2 Drosophila melanogaster cDNA clone SD02623 3prime similar to
 U21123: Drosophila melanogaster ena polypeptide gene, complete cds,
 mRNA sequence.
 ACCESSION AI531538
 VERSION AI531538.1 GI:4445673
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 460)
 AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
 Lewis,S. and Rubin,G.M.
 TITLE BCGP/HMMI Drosophila EST Project
 JOURNAL Unpublished (2001)
 COMMENT Contact: Stapleton, M.
 BCGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, estfruitfly@berkeley.edu
 Based upon the presence of a XhoI site followed by a run of 14 or
 more T residues at the beginning of the sequence, this clone
 probably contains an inverted insert. The resulting Poly-T sequence
 has been removed.

Plate: 26 row: B column: 11
 High quality sequence stop: 268.

FEATURES source

BASE COUNT 106 a 119 c 96 g 139 t
 ORIGIN
 Query Match 69.2%; Score 16.6; DB 10; Length 460;
 Best Local Similarity 82.6%; Pred. No.3.1e+02;
 Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 tcgtcgacgttcgttaacgttc 23
 ||||| ||||| ||||| ||||| |||||
 Db 80 TCGTCGATCGATCGTCAACATTC 102

RESULT 9

LOCUS AQ860867/c 464 bp DNA GSS 03-NOV-1999
 DEFINITION nbe00015L04r CUGI Rice BAC Library (EcoRI) Oryza sativa genomic
 clone nbe0015L04r, DNA sequence.
 ACCESSION AQ860867
 VERSION AQ860867.1 GI:6211324
 KEYWORDS GSS.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 464)
 AUTHORS Wing,R.A. and Dean,R.A.
 TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
 JOURNAL Unpublished (1998)
 COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: GGAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence start: 22
 High quality sequence stop: 363.
 Location/Qualifiers
 1..464
 /organism="Oryza sativa"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /clone="nbe0015L04r"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACindigo; Site_1: EcoRI; Site_2: EcoRI;
 Rice is the most important food crop in the world. Half of
 the world population, especially those inhabiting highly

FEATURES source

BASE COUNT 106 a 119 c 96 g 139 t
 ORIGIN
 Query Match 69.2%; Score 16.6; DB 10; Length 460;
 Best Local Similarity 82.6%; Pred. No.3.1e+02;
 Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;


```

Query Match      69.2%; Score 16.6; DB 10; Length 522;
Best Local Similarity 82.6%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 2 cgtcgacgttcgttaacgttcg 24
||||| ||| ||| ||| ||| ||| |||
Db 148 CGCCGAAGTTGGTTGACGTCG 126

RESULT 12
BG638934
LOCUS      BG638934      532 bp      mRNA      EST      23-APR-2001
DEFINITION LD38917.3prime LD Drosophila melanogaster embryo pOT2 Drosophila
melanogaster cDNA clone LD38917 3 similar to emb: Fban0015112
'signal transduction' located on: 2R 56B5-56B5;; 04/10/2001, mRNA
sequence.
ACCESSION  BG638934
VERSION     BG638934.1 GI:13769486
KEYWORDS    EST.
SOURCE      fruit fly.
ORGANISM    Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 532)
AUTHORS     Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
TITLE       BGP/HHMI Drosophila EST project
JOURNAL     Unpublished (2001)
COMMENT     Other ESTs: LD38917.5prime
Contact: Stapleton, M.
BGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST. est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or
more T residues at the beginning of the sequence, this clone was
polyadenylated. The resulting Poly-T sequence has been removed. hit
genomic AE003797: arm.2R [13841036..14128894]
estimated-cyto:55F6-56C1: 04/10/2001
Plate: LD.389 row: B column: 5
High quality sequence stop: 504.
FEATURES
source
1..532
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="LD38917"
/clone_lib="LD Drosophila melanogaster embryo pOT2"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="XLI Blue"
/notes="Organ: embryo; Vector: pOT2; Site_1: EcoRI; Site_2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2."
BASE COUNT      113 a 147 c 112 g 160 t
ORIGIN

Query Match      69.2%; Score 16.6; DB 11; Length 532;
Best Local Similarity 82.6%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 tcgtcgacgttcgttaacgttc 23
||||| ||| ||| ||| ||| ||| |||
Db 80 TCGTCGATCGATCGTCAACATTC 102

RESULT 13
A1976120
LOCUS      A1976120      537 bp      mRNA      EST      27-AUG-1999
DEFINITION EST270714 Schistosoma mansoni female, Phil Loverde/Joe Merrick
Schistosoma mansoni cDNA clone SMFB14 5' end, mRNA sequence.
ACCESSION  A1976120
VERSION     A1976120.1 GI:5789288
KEYWORDS    EST.
SOURCE      Schistosoma mansoni.
ORGANISM    Schistosoma mansoni.
Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
Rhabditophora; Eulcithophora; Revertospermata; Mediofusata;
Neodermata; Trematoda; Digenea; Strigeida; Schistosomatoidea;
Schistosomatidae; Schistosoma.
REFERENCE   1 (bases 1 to 537)
AUTHORS     Merrick,J.M., Osman,A., Loverde,P.T., Chandra,I., Glodek,A., Fraser
,C.M. and Lee,N.H.
TITLE       Generation of a Schistosoma sp. EST (SMEST) Catalog & Schistosome
Gene Index
JOURNAL     Unpublished (1998)
COMMENT     Contact: Norman H. Lee
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3529
Fax: 301 838 0208
Email: nhlee@tigr.org
Seq primer: M13 Reverse.
FEATURES
source
1..537
/organism="Schistosoma mansoni"
/db_xref="taxon:6183"
/clone="SMFB14"
/clone_lib="Schistosoma mansoni female, Phil Loverde/Joe
Merrick"
/sex="female"
/notes="Vector: pBluescript SK; Site_1: EcoRI; Site_2: XhoI
; directionally cloned cDNAs"
BASE COUNT      162 a 99 c 125 g 151 t
ORIGIN

Query Match      69.2%; Score 16.6; DB 10; Length 537;
Best Local Similarity 82.6%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 tcgtcgacgttcgttaacgttc 23
||||| ||| ||| ||| ||| ||| |||
Db 268 TCGTCAACGTTTGATACGTC 290

RESULT 14
BG795168/c
LOCUS      BG795168      552 bp      mRNA      EST      16-MAY-2001
DEFINITION UTSM SM256 UTSM Adult Mouse Skeletal Muscle Library Mus musculus
cDNA clone UTSM_SM256, mRNA sequence.
ACCESSION  BG795168
VERSION     BG795168.1 GI:14130738
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 552)
AUTHORS     Gallardo,T.D., Schageman,J.J., Pertsemildis,A., Garner,H.R.,
Williams,R.S. and Shohet,R.V.
UT Southwestern Medical Center, Adult Mouse Skeletal Muscle cDNA
Library
Unpublished (2001)
Contact: Schageman JJ
Shohet/Garner Labs
University of Texas Southwestern Medical Center
6000 Harry Hines Blvd., NA2.226, Dallas, TX 75390, USA
Tel: 214 648 1674
Email: Jeff.Schageman@UTSouthwestern.edu
cDNA library constructed by UTSM as a component of the Program for
Genomic Applications (PGA) and the Reynolds Heart Disease

```

...

...